NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

GLOBAL MODEL FORECASTS OF 2005 ATLANTIC TROPICAL CYCLONE FORMATIONS AFTER POST-PROCESSING TO ACCOUNT FOR INITIAL INTENSITY

by

Stephen G. Chesser

March 2008

Thesis Advisor:	Russell L. Elsberry
Second Reader:	Mark A. Boothe

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** March 2008	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Global Model Forecasts of 2005 Atlantic Tropical Cyclone Formations After Post-Processing to Account for Initial Intensity			**5. FUNDING NUMBERS**
6. AUTHOR(S) Stephen G. Chesser			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The objective of this thesis was to test the impact on Atlantic tropical cyclone formation forecasts during 2005 by three global models via a post-processing technique of adjusting the initial conditions to match the National Hurricane Center initial intensity information contained in the Combined Automated Response to Query (CARQ). Histograms of model analyses of the 850 – 500 mb relative vorticity and the 700 – 500 mb warm core, which are derived from the VORTRACK files, are created for CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt, and then are used to derive Lower Tercile Values (LTVs). These LTVs of relative vorticity and warm core for each model are used to adjust the initial conditions to agree with the CARQ intensity, and the LTV_{35} is considered to be the threshold intensity value for formation. These adjusted model forecasts are all superior to the unadjusted forecasts because many of the false alarms are eliminated. The adjusted model forecasts of relative vorticity and warm core are also converted into equivalent intensity forecasts, and a consensus of these intensities provides a useful indication of the evolution of an incipient tropical disturbance toward the tropical storm stage.

14. SUBJECT TERMS Tropical cyclone formation; Numerical model cyclone forecasts; Post-processing numerical models			**15. NUMBER OF PAGES** 125
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

GLOBAL MODEL FORECASTS OF 2005 ATLANTIC TROPICAL CYCLONE FORMATIONS AFTER POST-PROCESSING TO ACCOUNT FOR INITIAL INTENSITY

Stephen G. Chesser
Captain, United States Air Force
B.S., University of South Alabama, 1999

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
March 2008

Author: Stephen G. Chesser

Approved by: Russell L. Elsberry
Thesis Advisor

Mark A. Boothe
Second Reader

Philip A. Durkee
Chairman, Department of Meteorology

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The objective of this thesis was to test the impact on Atlantic tropical cyclone formation forecasts during 2005 by three global models via a post-processing technique of adjusting the initial conditions to match the National Hurricane Center initial intensity information contained in the Combined Automated Response to Query (CARQ). Histograms of model analyses of the 850 – 500 mb relative vorticity and the 700 – 500 mb warm core, which are derived from the VORTRACK files, are created for CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt, and then are used to derive Lower Tercile Values (LTVs). These LTVs of relative vorticity and warm core for each model are used to adjust the initial conditions to agree with the CARQ intensity, and the LTV_{35} is considered to be the threshold intensity value for formation. These adjusted model forecasts are all superior to the unadjusted forecasts because many of the false alarms are eliminated. The adjusted model forecasts of relative vorticity and warm core are also converted into equivalent intensity forecasts, and a consensus of these intensities provides a useful indication of the evolution of an incipient tropical disturbance toward the tropical storm stage.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

ACKNOWLEDGMENTS

Professor Russell Elsberry provided expert direction for the thesis as well as constant guidance. Professor Patrick Harr provided guidance for the VORTRACK system and is thanked for his efforts to provide data sets. Mary Jordan provided much needed MATLAB expertise that was pivotal to the initial sorting of the dataset. Mark Boothe provided much needed Microsoft Excel and FORTRAN expertise for dataset manipulation, for formatting, for charts and graphs and is thanked for serving as second reader. The VORTRACK system was sponsored by the National Oceanic and Atmospheric Administration, Joint Hurricane Testbed Project and the Office of Naval Research, Marine Meteorology Program. I would also like to thank my wife Jennifer for her continuous support throughout the completion of this degree.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Tropical cyclones have large impacts on the U. S. Air Force strategic planning and tactical operations, on long-range plans for and the delivery of materiel and personnel, and of course on air bases throughout the tropics and on the east coasts of continents. Fortunately, improvements have been made in operational global and regional models that have resulted in increased accuracy of 72-h tropical cyclone track forecasts, which has led in 2003 to an extension of preparation lead times when the National Hurricane Center (NHC) and the Joint Typhoon Warning Center (JTWC) began issuing track forecasts through 120 h.

Whereas the skill of tropical cyclone track has been dramatically improved during the past decade, the formation and intensification forecasts have not been similarly improved (Elsberry 2003). Forecasts of formation have become more important with the issuance of 120-h track forecasts since a tropical cyclone can form and intensify to hurricane stage during the 120-h forecast interval. For example, Hurricane Humberto was a short-lived tropical cyclone notable for its exceptionally rapid (within 19 hours) intensification near the coast of Texas from a mere tropical depression to a strong category 1 hurricane on the Saffir-Simpson Hurricane Scale (Blake 2007). This example of a tropical cyclone forming and intensifying into such a powerful storm within this 120 h window was all the more dangerous due to its formation in close proximity to the coastline. Thus, an urgent need exists for accurate prediction of tropical cyclone formation from operational models. From the U. S. military perspective, this is especially important because so many vulnerable military stations are either on the coastline or located within 50 miles of the coastline.

It is important for meteorologists to warn their customers with enough lead time to complete preparations. One approach to assist the forecasters in this warning task is to both establish the formation forecast accuracy of the global models and identify and understand factors that distinguish forecasts of vortices that are correctly forecast to intensify into tropical cyclones (developers) from forecasts of vortices that are forecast to

intensify into tropical cyclones, but actually do not (false alarms). These objectives are addressed in this thesis by examining forecasts of formation by three global numerical models.

A. NUMERICAL MODEL APPROACH TO TROPICAL CYCLONE FORMATION

1. Background

Generally global models have been examined for tropical cyclone formation forecasting since these models are integrated two to four times a day and predict the large-scale environmental conditions known to be important for tropical cyclone formation. The numerical model approach is an initial-value problem in which initial conditions at time zero (t_0) are defined and the model equations are solved in time increments (Δt) to predict the conditions at a later time (t_f). That is,

$$F\left(t_f\right) = F\left(t_0\right) + \sum S\left(t\right)\Delta t\,,$$

where $S(t)$ is the nonlinear evolution of the solution that is summed over the time increments between times t_0 and t_f.

In some model applications to tropical cyclone formation, the goal is to predict the time t_f at which the threshold condition(s) for a tropical cyclone [F(TC)] is first met or exceeded. For example, one of the threshold conditions from Cowan (2006) for declaring a formation in the GFS model is that the 850 - 500 mb average relative vorticity should exceed 3.75 x 10^{-5} s^{-1}. In the schematic of the predicted time evolution in Figure 1, the middle solution is considered to be the truth, so the correct tropical cyclone formation time is t_2. The upper and lower lines in Figure 1 represent model solutions that begin from larger or smaller F(t) values at time t_0 and thus would predict earlier (t_1) or later (t_3) formation times, respectively.

Three aspects of the use of global numerical models to predict tropical cyclone formation introduce uncertainty in the prediction problem, which is then more complicated than in the schematic in Figure 1. First, an uncertainty exists in the definition of the threshold value(s) F(TC) that define a tropical cyclone formation in

terms of model-predicted variables. Although the favorable environmental conditions for tropical cyclone formation are known, these environmental conditions are necessary but not sufficient conditions.

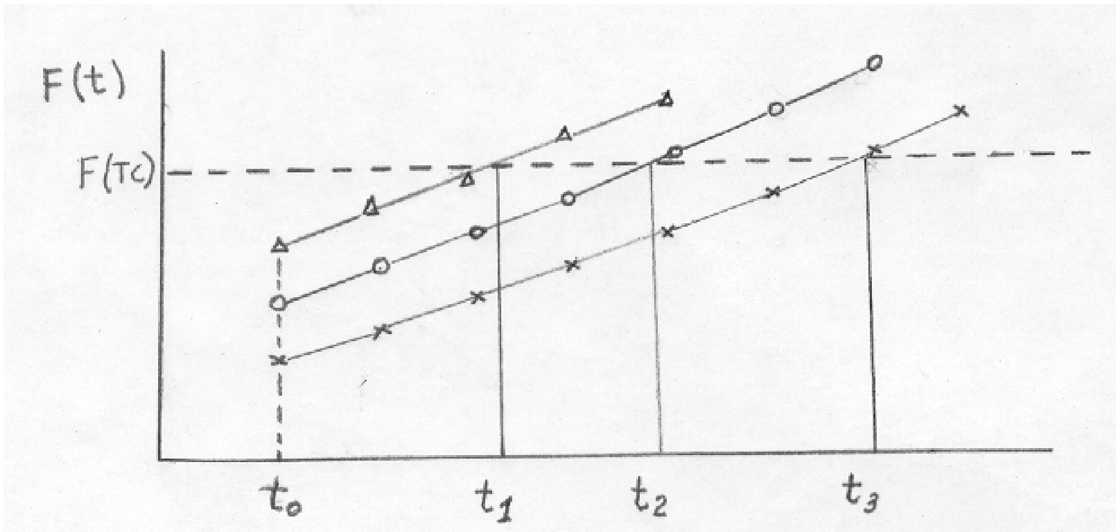


Figure 1 Time evolution schematic. Schematic of the time evolution of three numerical model solutions $F(t)$ beginning from time t_0, and the times (t_1, t_2, t_3) at which a threshold condition for tropical cyclone formation $F(TC)$ is first met. The middle solution (o symbols) with a formation time t_2 is considered to be the truth and the upper solution (Δ symbols) and lower solution (x symbols) begin from erroneous initial conditions and thus result in erroneous tropical cyclone formation times t_1 and t_3, respectively.

Second, uncertainty in the model initial conditions can lead to variability in the predicted formation time, as illustrated in Figure 1. Third, the numerical model solutions for the time tendencies in Equation (1) are not perfect and the errors generally increase with increasing forecast interval.

Each of these three contributions to uncertainty in using global models to predict tropical cyclone formation will be examined in the following subsections. The objectives are to understand the uncertainty in global model predictions of tropical cyclone formation, and to propose an alternate approach of post-processing the global model predictions to improve and extend to a longer time interval the formation predictions.

2. Definition of Formation

Cowan (2006) followed the practice of several previous studies in defining tropical cyclone formation as having occurred at the time of the first six-hourly entry in the Best Track of the National Hurricane Center (NHC) for Atlantic tropical cyclones. During the 2005 season studied by Cowan (2006), the initial intensity (maximum sustained surface wind speed anywhere in the cyclone) that corresponds to this first entry in the Best Track varies with each cyclone (Figure 2). Although the most common (mode) initial wind speed with 11 cases is 30 kt, ten cases had an initial wind speed of 25 kt (i.e., a tropical depression). However, one case had an initial speed of 35 kt (i.e., a tropical storm), and tropical cyclone 24L (Vince) had an initial speed of 55 kt. Since Cowan (2006) defined tropical cyclone formation in the GFS model predictions as the lower tercile value of the 850 - 500 mb relative vorticity (and of 700 - 500 mb warm core), she would have included model vortex relative vorticities corresponding to these times of initial entries in the NHC Best Track.

Although using the lower tercile values in the relative vorticities would exclude the 24L (Vince) case with an initial speed of 55 kt, the lower tercile value would likely come from some combination of the 11 (10) cases of 30 kt (25 kt) initial Best Track entries, and thus not correspond to the tropical depression (25 kt) or the tropical storm (34 kt) as a definition for formation.

The estimated intensity at the time of the initial Combined Automated Response to Query (CARQ) message for the 2005 Atlantic named tropical cyclones is also shown in Figure 2 (solid bars). This intensity in the CARQ message generated at the NHC is based primarily on satellite image interpretation. Notice that these first CARQ entry intensities range from: 15 kt (5 cases); 20 kt (13 cases); 25 kt (3 cases); 30 kt (1 case); and 40 kt (1 case), which is the 24L (Vince) case. Also shown in Figure 2 are the numbers of hours that CARQ messages exist prior to the first entry in the NHC Best Track for this set of 2005 Atlantic season tropical cyclones. Whereas for three tropical cyclones the CARQ existed for 36 h or less, in most cases the CARQs were available 42 h - 66 h before the initial entry in the NHC Best Track. In four cases, more than 100 h of CARQs were available prior to the start of the Best Track.

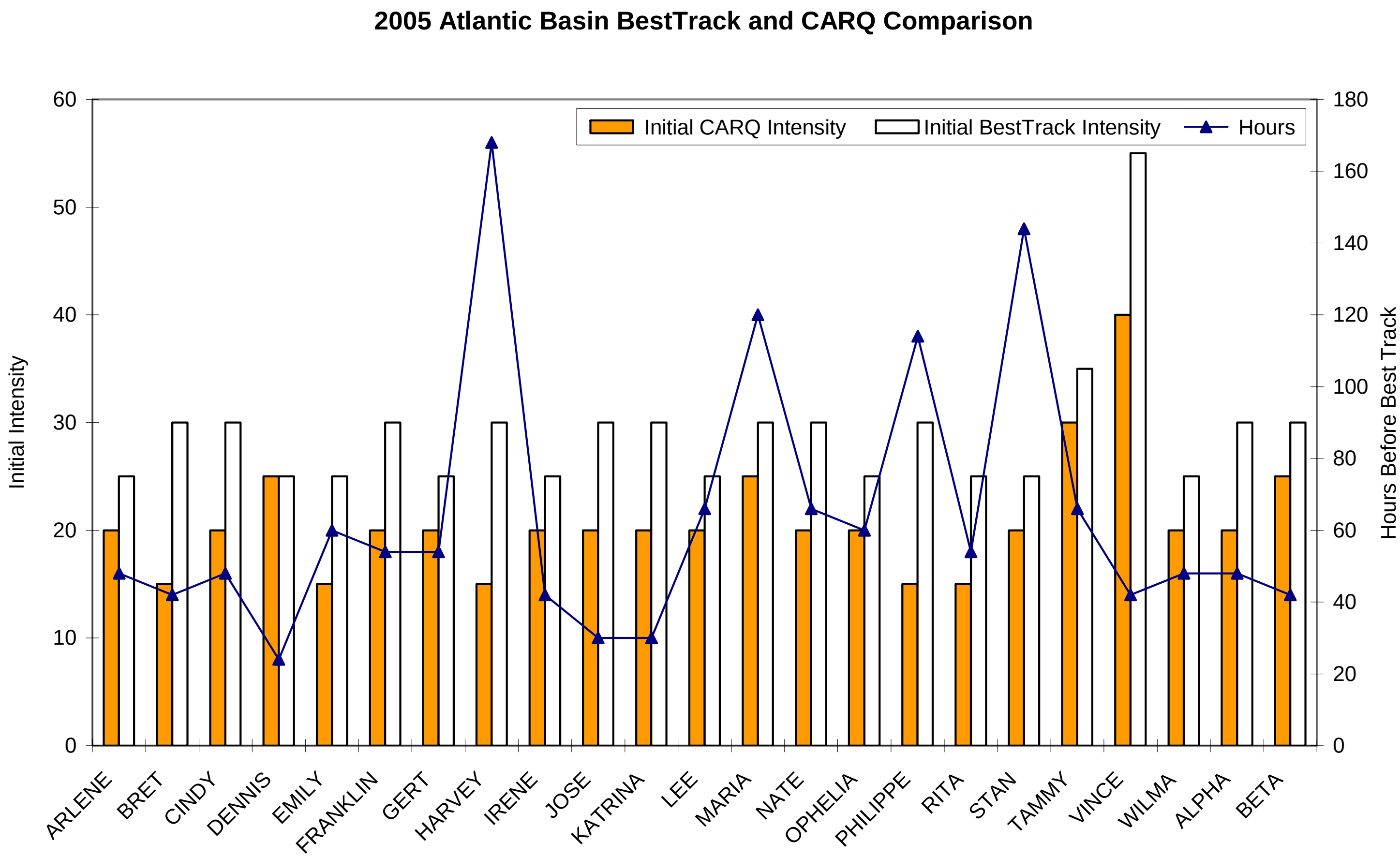


Figure 2 2005 Atlantic Basin Best Track and CARQ Comparison. Initial intensity (kt, left ordinate) of the tropical disturbances in the North Atlantic during the 2005 season from the initial CARQ (solid bars) and initial Best Track entry (open bars). The number of hours between the initial CARQ and the initial Best Track entry (right ordinate) is indicated by the solid line.

In contrast to later stages in which synthetic tropical cyclone observations (also called a bogus vortex) will be inserted to improve track predictions, sufficient observations are typically not available to ensure the model initial conditions have a vortex of the intensity (or the location) in the CARQ. Indeed, many of the Atlantic pre-tropical cyclone seedlings are in areas in which few in situ observations exist, so that satellite observations must be used to define the structure and location. Since the warm core temperature, within the 850 mb relative vorticity 1.5 x 10^{-5} s^{-1} isoline averaged from 700 - 500 mb relative to the environmental temperature may be only 0.2°C (Cowan 2006), satellite-derived temperature observations may not be adequate to define this structure in the global model conditions. In lieu of observations, the initial conditions may be dominated by the background field from the previous 6 h or 12 h forecast. For these reasons, the vortex intensity in the global model analyses and forecasts may depart significantly from the CARQ intensity. Warm core measurements are defined as a temperature difference between the vortex and the immediate environment (Harr 2006).

B. PLAN FOR THE THESIS

Even though uncertainty exists in the intensity estimates in the CARQ, and while these estimates are only discretized at 5 kt intervals, the hypothesis here is that these intensity estimates provide a more consistent indication of the formation timing than do the times of the first entry in the NHC Best Track. Thus, the global model analyses of 850 - 500 mb relative vorticities and 700 - 500 mb warm core magnitudes for the sample of 2005 Atlantic tropical cyclones will be compared with the CARQ intensities when these model variables are available from the VORTRACK (Harr 2006). Following Cowan (2006), the lower tercile values of these model variables at the various CARQ-based intensities will then be used to define the time of formation. Whereas some other fraction than one-third might be used (e.g., one-quarter or one-half), only the lower tercile definition of formation in the global models will be tested here.

The primary objective of this thesis is to test the impact on tropical cyclone formation forecasts of an improved specification of the initial conditions of three global models using the CARQ intensities as 'truth.' The CARQ data and the VORTRACK

files used to define the tropical cyclone formation-related variables are described in Chapter II. In Chapter III, the methodology for estimating the lower tercile values (LTV) for these variables that correspond to CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt will be presented. These LTVs are used to both adjust the initial values in the numerical models in the post-processing technique as in Figure 1 and serve as the ‘truth’ such that “formation” is declared when that LTV threshold is exceeded in the model. The tests of the initial condition post-processing technique for the three global models are summarized in Chapter IV. An alternate approach of converting the global forecasts of 850 - 500 mb relative vorticity and 700 - 500 mb warm core to the equivalent cyclone intensity is also introduced and tested in Chapter IV. Finally, the impacts on tropical cyclone formation forecasts by post-processing the initial conditions using CARQ intensities are summarized in Chapter V, and suggestions for future research to further improve the global model forecasts are also given.

THIS PAGE INTENTIONALLY LEFT BLANK

II. CARQ AND VORTRACK DATASETS

The CARQ data sets were downloaded from the Tropical Prediction Center (TPC) of the National Centers for Environmental Prediction (NCEP) ftp website ftp:tpc.ncep.noaa.gov for all invest regions, tropical disturbances, tropical depressions, tropical storms, and hurricanes during 2005. A script MATLAB code (originally developed by Mary Jordan of the Department of Meteorology) was used to initially sort the raw data files into the desired categories of CARQ files for each disturbance (Table 1). These CARQ files were extracted from the time of the initial CARQ until the disturbance surpassed tropical storm strength (34 kt) to ensure representative Lower Tercile Values (LTV) at 20, 25, 30, and 35 kt thresholds.

It was also noted that some extended periods of CARQs existed when no VORTRACK values were available. According to Professor Patrick Harr, these early in the storm sequences may correspond to periods when vortices were being tracked by the VORTRACK algorithms, but the 850 mb relative vorticity was less than the 1×10^{-5} s^{-1} threshold required for matching with one of the Atlantic tropical cyclones. Although it might have been possible to expand the sample sizes by checking for these weaker vortices that correspond to the CARQ sequences, time was not available for such a thorough search of the VORTRACK files.

An example of the additional information that is available from the CARQ file for Tropical Storm Harvey (08L) before the initial time in the Best Track file is given in Table 1. Notice that the initial intensity is 15 kt and 168 hours (28 additional 6-hour periods) of CARQ intensity information are available before the initial time in the Best Track file that Cowan (2006) used as a starting point.

Having determined the Date-Time-Group (DTG) with CARQ intensities in this range and the beginning of each of the 2005 Atlantic season tropical cyclones in Figure 2, the corresponding 850 mb relative vorticity, 850 - 500 mb relative vorticity, and 700 - 500 mb warm core values from the VORTRACK (provided by Professor Patrick Harr) were extracted. When some of the VORTRACK values are missing (usually at 06 UTC

or 18 UTC) and CARQ intensities are available, linear interpolation is used to fill in the missing value to maximize the sample size.

Table 1. Summary of data relating to 2005 Atlantic season. List of Atlantic tropical cyclones by name and number used in this thesis. Initial intensities (kt) in the CARQ and Best Track are compared in the third and fourth columns. The number of hours that a CARQ intensity value was available prior to the initial time a Best Track intensity value was available is given in the last column.

Name	Number	Initial CARQ Intensity	Initial BestTrack Intensity	Hours of CARQ
ARLENE	01L	20	25	48
BRET	02L	15	30	42
CINDY	03L	20	30	48
DENNIS	04L	25	25	24
EMILY	05L	15	25	60
FRANKLIN	06L	20	30	54
GERT	07L	20	25	54
HARVEY	08L	15	30	168
IRENE	09L	20	25	42
JOSE	11L	20	30	30
KATRINA	12L	20	30	30
LEE	13L	20	25	66
MARIA	14L	25	30	120
NATE	15L	20	30	66
OPHELIA	16L	20	25	60
PHILIPPE	17L	15	30	114
RITA	18L	15	25	54
STAN	20L	20	25	144
TAMMY	22L	30	35	66
VINCE	24L	40	55	42
WILMA	25L	20	25	48
ALPHA	26L	20	30	48
BETA	27L	25	30	42

A total of 391 matches were found between the CARQ intensity, DTGs and the corresponding VORTRACK values from the GFS, NOGAPS, and UKMO analyses. The GFS had the most matches with 134 cases, while NOGAPS was next with 131 cases, and UKMO model had the fewest matches with only 126 cases.

III. METHODOLOGY

A. METHODOLOGY FOR CARQ-BASED INTENSITY LOWER TERCILE VALUES

1. Defining Formation Using Lower Tercile Values (LTVs)

The objective of this section is to estimate a set of consistent Lower Tercile Values (LTVs) of 850 - 500 mb relative vorticity and 700 - 500 mb warm core values that correspond to the CARQ intensities. These values for relative vorticity and warm core are to be used as the threshold values [F(TC) in Figure 1] that correspond to tropical cyclone formation in each of the three models depending on whether the CARQ intensity is 25 kt or 35 kt (or perhaps 30 kt). It is clear from the histograms in Appendices A, B, and C of 850 - 500 mb relative vorticity and 700 - 500 mb warm core that different thresholds apply in the NOGAPS, GFS, and UKMO models because of differences in data assimilation, horizontal grid spacing, model physics, and other factors. In addition, these LTVs are also to be used in the post-processing technique to adjust the initial values of 850 - 500 mb relative vorticity and 700 - 500 mb warm core to correspond to the CARQ intensity at the initial time as in Figure 1. Thus, it will be assumed these initial values evolve smoothly with increasing intensity.

Based on the histograms of the 850 - 500 mb relative vorticity for each model as presented in the Appendices, a summary as a function of the CARQ intensities is given in Table 2. As is evident in these histograms, relatively small numbers of cases are available for each CARQ intensity category for each model in Table 2. It is also clear that overlaps of relative vorticity values exist as the CARQ intensities increase from 20 kt to 35 kt. These ranges indicate that the model initial relative vortices are not constrained to agree with the CARQ intensity by use of a bogus vortex, or synthetic tropical cyclone observations as are imposed in the NOGAPS and UKMO models when the tropical

cyclone is named (CARQ intensity of 35 kt). To apply the post-processing technique, a single value of 850 - 500 mb relative vorticity is required for each CARQ intensity category.

Table 2. Summary of the 850 - 500 mb relative vorticity (10^{-5} s^{-1}) statistics for the NOGAPS, GFS, and UKMO models as a function of the CARQ intensities for selected developer cases during the 2005 Atlantic hurricane season. The range and average, and the Lower Tercile Values (LTV) determined by a statistical fit, by a counting of the cases technique, and the final values after smoothing are provided.

	20 kt	25 kt	30 kt	35 kt
Model		**NOGAPS**		
Cases	27	39	31	21
Range	1.22 to 3.93	1.85 to 5.69	2.28 to 6.67	3.27 to 6.65
Avg	2.64	3.44	4.64	5.65
LTV Stat	2.12	3.13	3.74	4.45
LTV Case	2.47	3.09	3.94	5.63
LTV Final	2.32	3.03	3.74	4.45
Model		**GFS**		
Cases	20	47	32	20
Range	1.68 to 4.65	2.53 to 6.88	2.24 to 10.28	4.00 to 6.89
Avg	3.37	4.04	4.64	5.08
LTV Stat	2.67	3.98	4.92	4.96
LTV Case	3.05	3.65	4.07	4.53
LTV Final	3.15	3.61	4.07	4.53
Model		**UKMO**		
Cases	25	43	27	19
Range	0.99 to 5.58	1.15 to 6.55	2.38 to 7.41	3.73 to 7.74
Avg	2.86	3.92	4.89	5.75
LTV Stat	2.52	2.95	4.06	5.07
LTV Case	2.09	3.17	4.37	5.53
LTV Final	2.09	3.17	4.37	5.53

One option would be to assign the average 850 - 500 mb relative vorticity as the single value for each intensity category and each model as listed in Table 2. Notice that these average values are not in the middle of the ranges because the distributions are not Gaussian, but they are generally skewed towards larger values (see model histograms in Appendices). A median value (not shown) would be another option, so that 50% of the

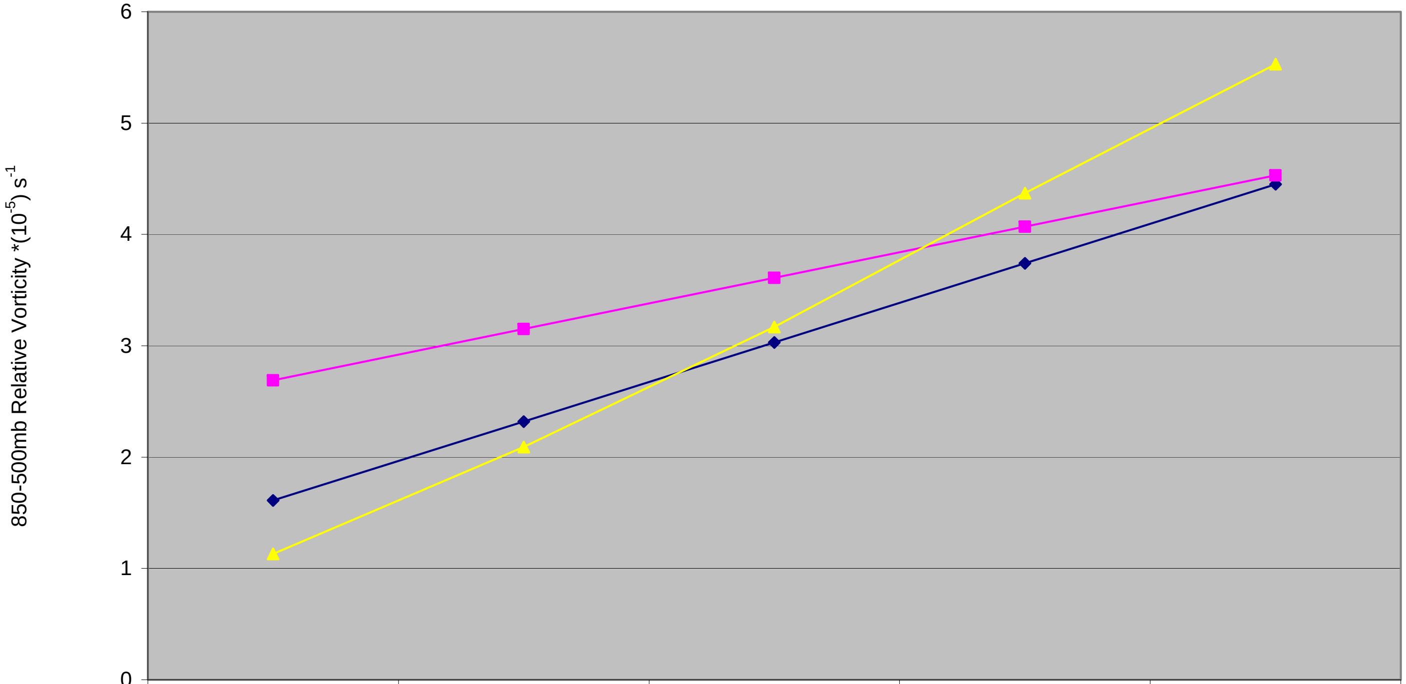
sample would be larger. Rather, to be consistent with using the LTV as the threshold value for tropical cyclone formation of the 850 - 500 mb relative vorticity, the single value in the post-processing technique (Figure 1) must be the LTV for the initial CARQ intensity, and using the LTV is consistent with Cowan (2006) as well.

Two methods of estimating the LTV were examined since the sample sizes are small. The statistical fit technique presumes a Gaussian distribution, which does not apply for many of the 850 - 500 mb relative vorticity histograms. The second case method is performed by a simple counting of the number of cases and interpolating the relative vorticity corresponding to the lower tercile. As indicated for each CARQ intensity and model in Table 2, the statistical and case LTVs may provide some quite different LTVs of 850 - 500 mb relative vorticities. For example, the LTVs for NOGAPS model corresponding to a CARQ intensity of 35 kt are 4.45 x 10^{-5} s^{-1} from the statistical fit and 5.63 x 10^{-5} s^{-1} via the case method. Likewise, the 0.85 x 10^{-5} s^{-1} difference between the statistical fit and the case method for the GFS model for 30 kt CARQ intensities may be attributed to the uncertainty arising from small samples.

Considering the dual purpose of using the LTV as the initial values of the 850 - 500 mb relative vorticity as the threshold and as the initial values in the post-processing technique, the desire is to have a consistent variation with increasing CARQ intensity up to the 35 kt threshold. The LTV Final entries in Table 2 are smoothed 850 - 500 mb relative vorticities that vary almost linearly between 20 and 35 kt, and they are intended to represent the evolution of model vortices during the early stages of a tropical cyclone.

The variation of the smoothed LTV Final 850 - 500 mb relative vorticities with CARQ intensities are depicted in Figure 3. Although NOGAPS and GFS LTV Final values at a CARQ intensity of 35 kt are nearly identical (Table 2), the slopes of the LTV lines are different. The steepest increase of relative vorticity with increasing CARQ intensities is found in the UKMO model, which has the smallest LTV at 20 kt and largest LTV at 35 kt. Notice in Figure 3 that the LTV Final values for each model have been extrapolated down to a CARQ intensity of 15 kt. Since such a CARQ intensity is rarely specified, an insufficient number of cases is available to calculate the LTVs so that a consistent initial value of 850 - 500 mb relative vorticity would be available for the post-processing step (Figure 1) even if the intensity was only 15 kt.

Model Finalized Lower Tercile Values (LTV) for 850-500 mb Relative Vorticity



	15	20	25	30	35
NGP	1.61	2.32	3.03	3.74	4.45
GFS	2.69	3.15	3.61	4.07	4.53
UKM	1.13	2.09	3.17	4.37	5.53

Wind (Knots)

Figure 3 Smoothed Lower Tercile Values (LTVs) for 850 - 500 mb relative vorticity (10^{-5} s^{-1}) corresponding to the CARQ intensities (kt) for the NOGAPS (diamond symbols), GFS (square symbols), and the UKMO (triangle symbols).

Table 3. Summary of the 700 - 500 mb warm core (°C) statistics for the NOGAPS, GFS, and UKMO models as a function of the CARQ intensities for selected developer cases during the 2005 Atlantic hurricane season. The number of cases, the range and average, and the Lower Tercile Value (LTV) determined by a statistical fit, by a counting of the cases technique, and final values after smoothing are provided.

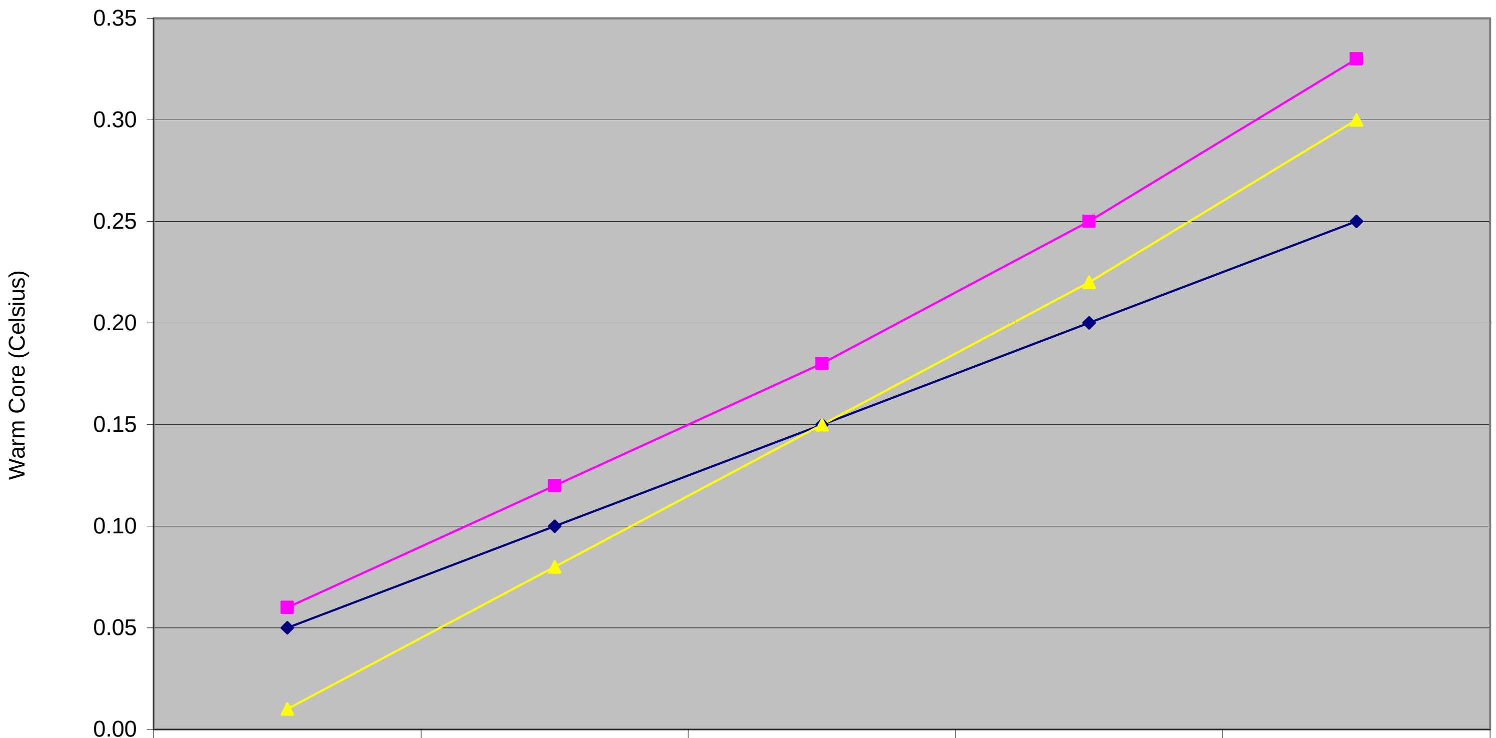
	20 kt	25 kt	30 kt	35 kt
Model		**NOGAPS**		
Cases	27	39	31	21
Range	0.58 to -0.14	0.69 to -0.08	0.75 to -0.06	0.81 to -0.01
Avg	0.19	0.22	0.33	0.36
LTV Stat	0.10	0.18	0.21	0.26
LTV Case	0.09	0.13	0.25	0.26
LTV Final	0.10	0.15	0.20	0.25
Model		**GFS**		
Cases	20	47	32	20
Range	0.50 to 0.02	0.57 to -0.09	0.82 to 0.02	0.08 to 0.05
Avg	0.22	0.25	0.36	0.37
LTV Stat	0.18	0.13	0.29	0.33
LTV Case	0.16	0.16	0.22	0.23
LTV Final	0.12	0.18	0.25	0.33
Model		**UKMO**		
Cases	25	43	27	19
Range	0.59 to -0.18	0.75 to -0.16	1.01 to -0.17	1.10 to 0.03
Avg	0.13	0.32	0.43	0.44
LTV Stat	0.08	0.14	0.22	0.39
LTV Case	0.06	0.26	0.32	0.32
LTV Final	0.08	0.15	0.22	0.30

The corresponding summary of 700 - 500 mb warm core statistics for the three models and CARQ intensities ranging from 20 kt to 35 kt is provided in Table 3. The numbers of cases for each model and intensity are the same as in Table 1. Notice that the ranges of 700 - 500 mb warm cores actually include negative values in the NOGAPS model even at a CARQ intensity of 35 kt, and in the UKMO model up to intensities of 30 kt. Although the averages of the warm cores for the three models are not that different, the average warm core trends with increasing CARQ intensities are somewhat different. Both the GFS and UKMO models have average warm core differences of only 0.01°C between 30 kt and 35 kt.

As was done for the 850 - 500 mb relative vorticities (Table 2), the LTVs of the 700 - 500 mb warm cores were derived by the statistical fit and the case method, and then a smoothed LTV evolution with CARQ intensity was derived (Table 3). Perhaps because the 700 - 500 mb warm core histogram distributions tended toward Gaussian more so than the 850 - 500 mb relative vorticity histograms, the statistical fit approach provided a satisfactory evolution with CARQ intensity. Thus, the LTV Final 700 - 500 mb warm core values were more similar to the statistical fit values than to the case method values. These LTV Final warm core values only vary from about 0.10°C for a CARQ intensity of 20 kt to about 0.30°C for an intensity of 35 kt (Table 3).

These LTV Final 700-500 warm core values vary smoothly (Figure 4) with increasing CARQ intensity from 20 kt to 35 kt. As was the case for the 850 - 500 mb relative vorticities (Figure 3), the UKMO model has the largest slope with increasing intensity, although the differences are relatively small. These smoothed LTVs for the 700 - 500 mb warm core were also extrapolated backwards to a CARQ intensity of 15 kt so such values are available for the post-processing technique in the rare cases that such a unique CARQ intensity is specified.

Model Finalized Lower Tercile Values (LTV) for 700-500 mb Warm Core



	15	20	25	30	35
NGP	0.05	0.10	0.15	0.20	0.25
GFS	0.06	0.12	0.18	0.25	0.33
UKM	0.01	0.08	0.15	0.22	0.30

Wind (Knots)

Figure 4 As in Figure 3, except for 700 - 500 mb warm core (°C).

In summary, smoothly varying LTVs have been defined for both 850 - 500 mb relative vorticities and the 700 - 500 mb warm cores for the NOGAPS, GFS, and UKMO models based on the VORTRACK files from the developing cases during the 2005 Atlantic hurricane season. These LTVs can be used as threshold values [F(TC) in Figure 1] in the model forecasts of 850 - 500 mb relative vorticity and 700 - 500 mb warm core if tropical cyclone formation is defined as 25 kt or 35 kt (or even 30 kt, if desired). In the post-processing technique, the initial values in the models can be adjusted to agree with these LTVs that correspond to the CARQ intensity from the NHC based primarily on the satellite imagery interpretation. Such an adjustment is proposed to improve both the timing of the formation and to reduce the numbers of false alarms by the NOGAPS, GFS, and UKMO models, provided the CARQ intensities are representative of the actual conditions in the atmosphere and the model forecasts of the 850 - 500 mb relative vorticity and 700 - 500 mb warm core tendencies are accurate. That is, the assumption in the post-processing technique in Figure 1 is that the primary uncertainty in the time at which F(TC) will be exceeded is associated with improper initial conditions in the model, and that these initial values can be corrected using the CARQ intensities.

2. Formation Timing Uncertainty Due to Initial Condition Variability

As indicated in Figure 1, variability in the initial conditions of either the 850 - 500 mb relative vorticity or the 700 - 500 mb warm core variables to define tropical cyclone formation with similar model-depicted time tendencies can lead to significant differences in the time that the threshold value is exceeded. Evidence that the initial condition variability may be a major factor in the formation timing uncertainty is in Figure 5 (Figure 4, Cowan 2006). After only 12 h GFS forecasts of the 850 - 500 mb relative vorticity and 700 - 500 mb warm core, a large fraction of the developers are in the proper upper-right quadrant in which the formation thresholds for both variables are exceeded. However, other developers are not in the correct quadrant:

i) Two in the upper-left quadrant meet the warm core threshold, but have slightly small vorticities;

ii) Three in the lower-left quadrant have too small warm cores and too small vorticities; and

iii) Two in the lower-right quadrant have failed because of the too-small warm core values.

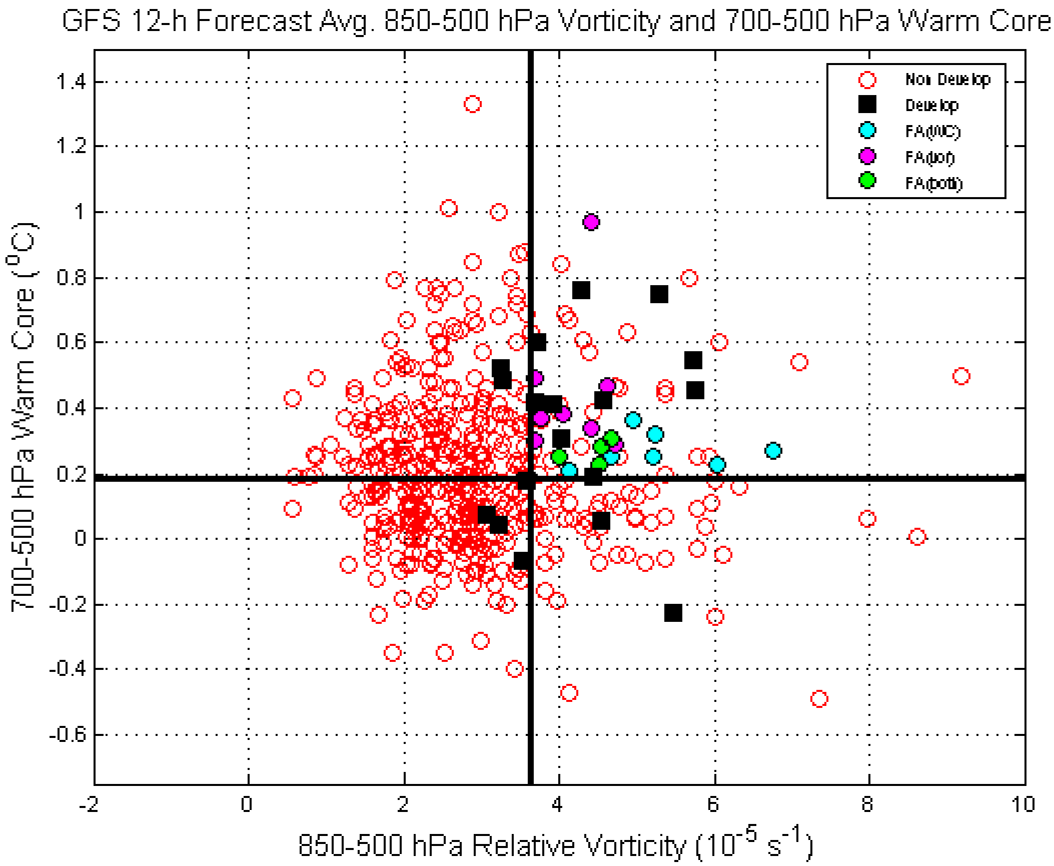


Figure 5 Scatterplot of 850 - 500 mb average relative vorticity and 700 - 500 mb warm core for the 12 h GFS vortices in the VORTRACK data base of 2005 Atlantic season tropical cyclones from Cowan (2006; Figure 4). The dark lines are the threshold values for these two variables that if surpassed would indicate tropical cyclone formation. The symbols (see inset) represent non-developers (open circles), developers (squares), and three kinds of False Alarms (filled circles).

Since these failures to predict developers in the proper quadrant in Figure 5 occur after only 12 h, it is hypothesized that these errors can be primarily attributed to improper initial conditions for these variables. As shown in the histograms for the GFS model in Appendix B (850 - 500 mb relative vorticity in Figures B-1 to B-4 and 700 - 500 mb warm core in Figures B-5 to B-8), some excessive values of 850 - 500 mb relative vorticity and 700 - 500 mb warm core exist at the various CARQ intensity categories. It

is hypothesized that a post-processing step in which the lower threshold values are used to raise (lower) insufficient (excessive) values would improve the predictions of formation timing.

3. Use of LTVs as 'Truth'

Having calculated the Lower Tercile Values (LTV) of both the 850 - 500 mb relative vorticity (Figure 3) and the 700 - 500 mb warm core (Figure 4) at 20 kt, 25 kt, 30 kt, and 35 kt, the tropical cyclone formation can then be defined as the time that the tropical disturbance reached 25 kt according to the CARQ data. For example, tropical cyclone formation as a 25 kt storm in the NOGAPS model would be defined when the 850 - 500 mb relative vorticity exceeded 3.03×10^{-5} m s^{-1} (Table 2) and the 700 - 500 mb warm core exceeded 0.15°C (Table 3). For the NCEP GFS (UKMO) model, the corresponding values that would define formation as 25 kt would be 3.61×10^{-5} m s^{-1} (3.17×10^{-5} m s^{-1}) and 0.18°C (0.15°C), respectively. By using these combinations of LTVs as the thresholds to define the formation time in each of the models, they become the 'truth' for evaluating the model forecasts.

Other alternatives would be to define a formation time as 35 kt (or perhaps 30 kt) to explore whether the models may have more skill in forecasting the transition from a tropical depression to a tropical storm. This transition is also an important forecast challenge as many of the Atlantic developers in this sample remained at the tropical depression stage for some time before becoming better organized, and once they reached tropical storm stage they continued to develop.

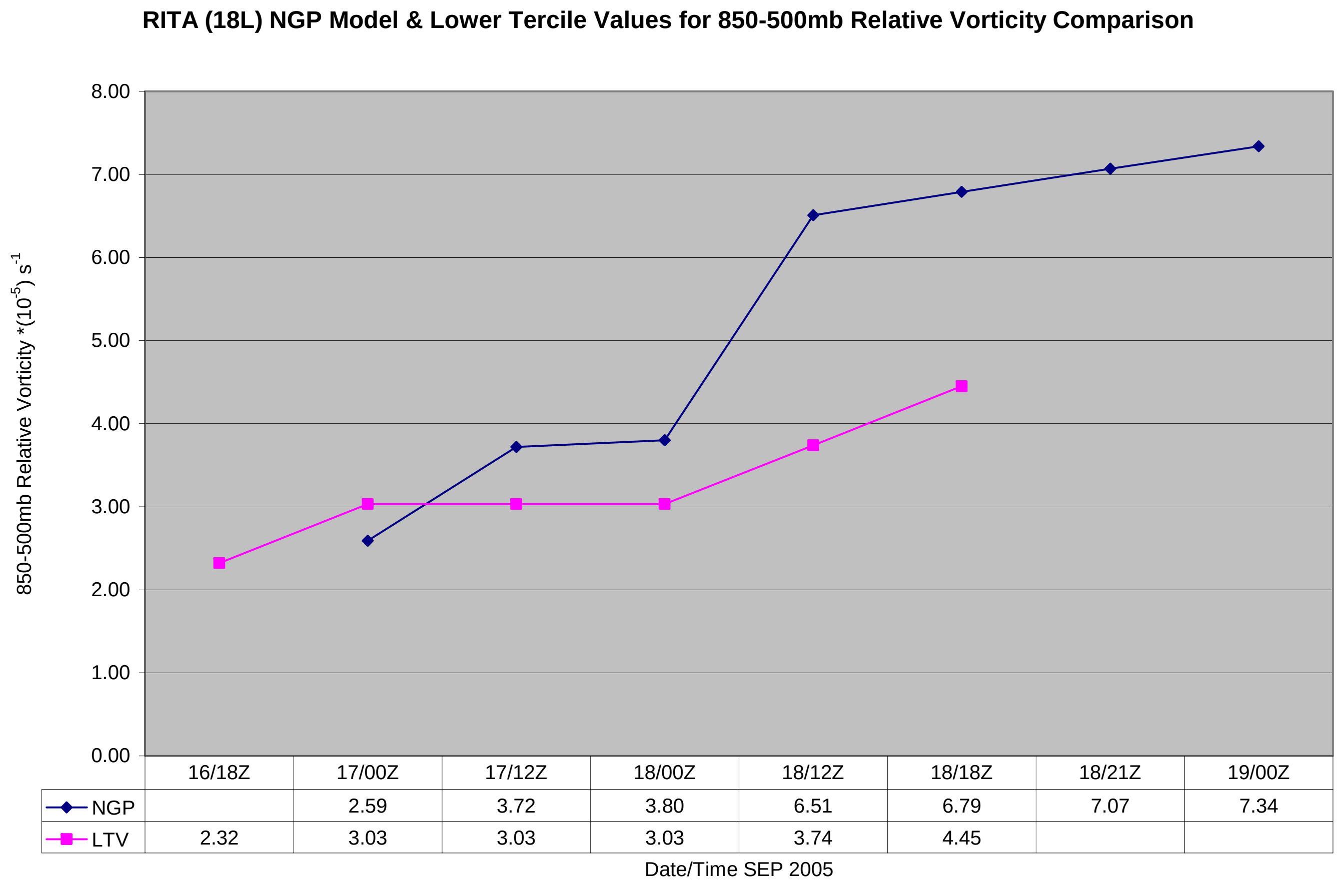
Thus, the test for each of the models is given that the initial CARQ intensity is 20 kt, 25 kt, or 30 kt, what is the skill of the model in forecasting the time at which the two LTV thresholds for that model are exceeded? These LTV threshold values correspond to 35 kt for the NOGAPS model are 4.45×10^{-5} m s^{-1} (Table 2) and 0.25°C (Table 3). For the GFS (UKMO), these LTV thresholds for 35 kt are 4.53×10^{-5} m s^{-1} (5.53×10^{-5} m s^{-1}) and 0.33°C (0.30°C).

To illustrate the effect of using the LTVs corresponding to the 35 kt CARQ intensity values as the 'truth' in re-defining the initial conditions, and for when the

threshold for 35 kt is exceeded, consider the case of TC Rita (18L). The LTVs of the 850 - 500 mb relative vorticity (Figure 6) and 700 - 500 mb warm core (Figure 7) that correspond to the CARQ values, which are considered to be the truth, have discrete values that correspond to 20 kt, 25 kt, 30 kt, and 35 kt. These truth values increase somewhat regularly until the values corresponding to 35 kt are achieved. However, the NOGAPS 850 - 500 mb relative vorticity (Figure 6) and 700 - 500 mb warm core values (Figure 7) from the VORTRACK for this TC do not correspond well to these truth values since no synthetic observations or bogus vortex appropriate for the CARQ intensity are inserted in the NOGAPS initial conditions. For most of the tropical cyclones in this sample, the NOGAPS relative vorticities are generally too low, so that even a perfect forecast of the time tendency of relative vorticity from these initial conditions might be expected to predict that the LTV threshold of relative vorticity for 35 kt would be achieved too late.

Similarly, the initial 700 - 500 mb warm core values for NOGAPS from the VORTRACK do not correspond well with the truth values (Figure 7). Even a perfect forecast of the time tendency of the warm core would be expected to result in a late forecast that the LTV corresponding to 35 kt would be exceeded when the initial values are too low as in Figure 7.

The next step is then to post-process the NOGAPS forecast by adjusting the initial conditions from the VORTRACK for the NOGAPS model to agree with the truth values of 850 - 500 mb relative vorticity and 700 - 500 mb warm core from the CARQ intensity at the initial time. The objective is to test whether this initial conditions adjustment, while maintaining the same model-predicted tendencies of relative vorticity and warm core, will produce more accurate forecasts of the times that the thresholds corresponding to 35 kt intensities will be exceeded. The metric will be the number of hours that the adjusted initial condition forecast is closer to the actual time that the initial 35 kt intensity occurred compared to the un-adjusted initial condition for NOGAPS. For this 35 kt test, all cases in the development sample from the VORTRACK outputs of the three models that also have a CARQ with an intensity less than 35 kt will be included. Similarly, tests could also be made for 30 kt and for 25 kt to determine model accuracy for these intensities, but progressively fewer forecasts would be available for evaluation.



	16/18Z	17/00Z	17/12Z	18/00Z	18/12Z	18/18Z	18/21Z	19/00Z
NGP		2.59	3.72	3.80	6.51	6.79	7.07	7.34
LTV	2.32	3.03	3.03	3.03	3.74	4.45		

Figure 6 Time evolution of 850 - 500 mb relative vorticity (10^{-5} s^{-1}) values corresponding to the CARQ intensities (squares) versus from NOGAPS initial analyses (diamonds) for TC Rita (18L) from 1800 UTC 16 September 2005.

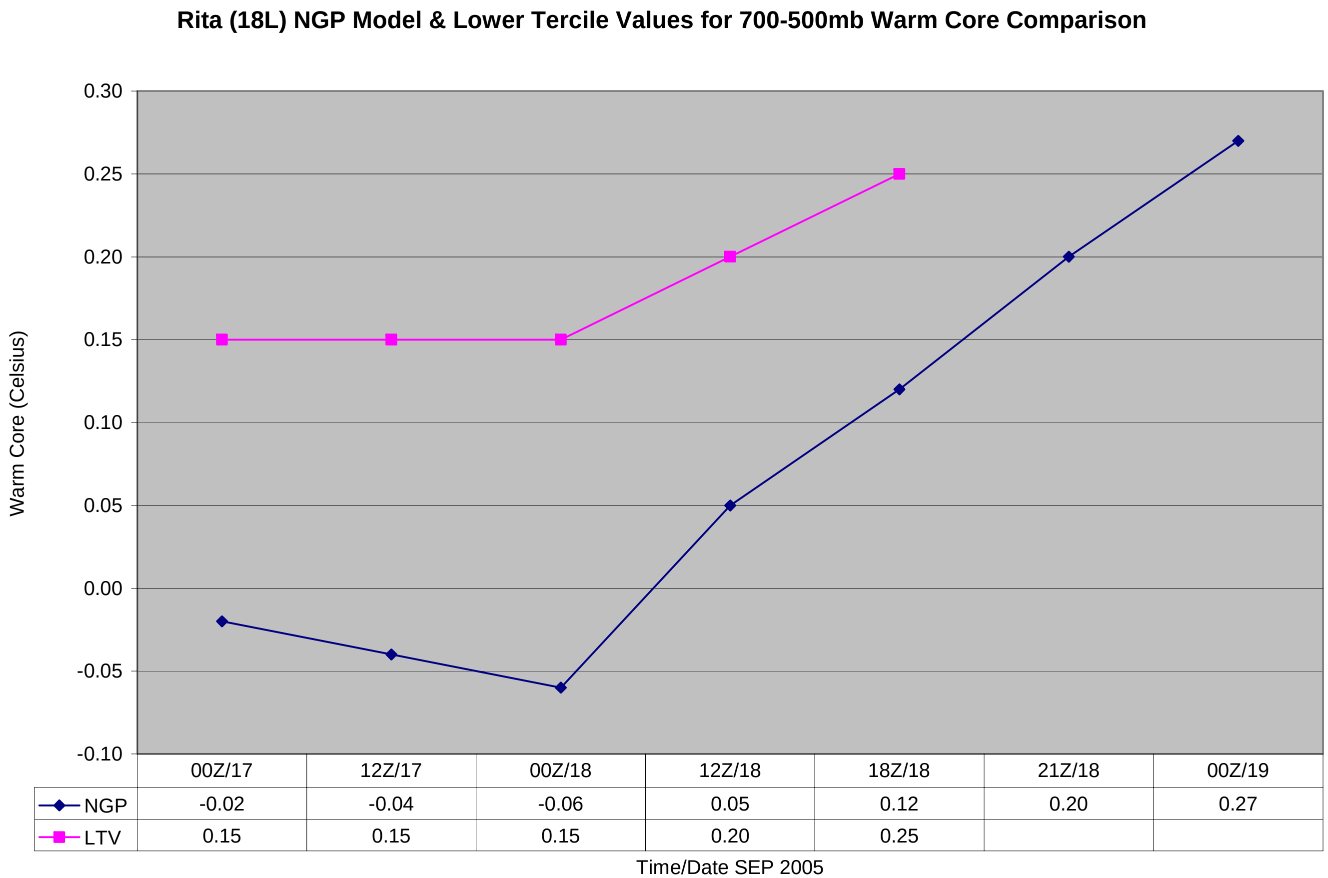


Figure 7 As in Figure 6, except for time evolution of 700 - 500 mb warm core (°C) for TC Rita (18L) from 1800 UTC 16 September 2005.

The GFS and UKMO models will be similarly evaluated for their skill in predicting 35 kt intensities with appropriately adjusted initial conditions and the relative vorticities and warm core thresholds appropriate for those intensities in Table 2 and Table 3. In addition, a consensus of the three model forecast times at which the 25 kt, 30 kt, and 35 kt intensities will occur might be calculated. Typically, a consensus forecast will be superior to the individual model forecasts over a sufficient sample size.

B. APPLICATION OF THE POST-PROCESSING TECHNIQUE

An example of the application of the post-processing technique is given in Table 4 for the NOGAPS analysis and 120 h forecast of 850 - 500 mb relative vorticity and 700 - 500 mb warm core. In this case from the first named storm of the 2005 season (01L/Arlene), the NOGAPS initial analysis had an 850 - 500 mb relative vorticity of $3.241 \times 10^{-5}\ s^{-1}$. Since the LTV Final relative vorticity in Table 2 for a CARQ intensity of 20 kt is $2.32 \times 10^{-5}\ s^{-1}$, a relative vorticity correction of $-0.921 \times 10^{-5}\ s^{-1}$ is required so that the adjusted relative vorticity value after the post-processing step as in Figure 1 begins with the LTV Final value. In this case, the NOGAPS forecasts the relative vorticity will decrease in time through the 120 h period. Consequently, the adjusted relative vorticity also decreases in time. The NOGAPS initial analysis had a warm core of -0.065°C. Since the NOGAPS LTV Final warm core value in Table 3 for a CARQ intensity of 20 kt is 0.10°C, a warm core correction of 0.165°C is applied in the post-processing step. Notice that the NOGAPS forecast of the 700 - 500 mb warm core is rather erratic with a rapid increase to 0.567°C after only 12 h and then a decrease to -0.039°C at 24 h. This diurnal-type variability continues for several days in this forecast. Applying the positive correction to all forecast values does not change the diurnal variability, and all of the adjusted warm core values are positive.

Table 4. Example of the post-processing technique as in Figure 1 at 0000 UTC 8 June 2005 for a tropical disturbance that later became 01L (Arlene). Whereas the NOGAPS initial analysis had an 850 - 500 mb relative vorticity of $3.241\text{x}10^{-5}$ s^{-1}, the Lower Tercile Value (LTV) for the CARQ intensity of 20 kt is $2.32\text{x}10^{-5}$ s^{-1}, so a correction of -0.921 x 10^{-5} s^{-1} is applied to obtain the Adjusted relative vorticity. Similarly, the initial 700 - 500 mb warm core (WC) value and the LTV of 0.10°C corresponding to a CARQ intensity of 20 kt leads to a WC correction of 0.165°C to obtain the Adjusted WC equal to the LTV.

Forecast Time (h)	NOGAPS Vorticity	Vorticity Correction	Adjusted Vorticity	NOGAPS Warm Core	Warm Core Correction	Adjusted Warm Core
0	3.241	-0.921	2.320	-0.065	0.165	0.100
6	2.875	-0.921	1.954	0.230	0.165	0.395
12	2.554	-0.921	1.633	0.567	0.165	0.732
18	2.535	-0.921	1.614	0.184	0.165	0.349
24	2.494	-0.921	1.573	-0.039	0.165	0.126
30	2.251	-0.921	1.330	0.177	0.165	0.342
36	2.395	-0.921	1.474	0.431	0.165	0.596
42	2.209	-0.921	1.288	0.160	0.165	0.325
48	2.274	-0.921	1.353	0.411	0.165	0.576
60	2.246	-0.921	1.325	0.156	0.165	0.321
72	2.176	-0.921	1.255	0.533	0.165	0.698
84	2.178	-0.921	1.257	0.578	0.165	0.743
96	2.298	-0.921	1.377	0.421	0.165	0.586
108	1.933	-0.921	1.012	0.373	0.165	0.538
120	2.019	-0.921	1.098	0.348	0.165	0.513

Using a program developed by Mark Boothe, the post-processing step for the 850 - 500 mb relative vorticity and 700 - 500 mb warm core analyses and forecasts was applied for the NOGAPS, GFS, and UKMO models for the same 2005 Atlantic season tropical storms in the developmental sample. Results of the post-processing technique for each of these models will be presented in the following sections.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. MODEL FORECASTS AFTER POST-PROCESSING

A. NOGAPS FORECASTS AFTER POST-PROCESSING

As indicated in Figure 1, the objective in the post-processing technique of adjusting the initial values of the 850 - 500 mb relative vorticity and the 700 - 500 mb warm core values in the model is to improve the forecasts of the times at which the threshold values (here LTVs) corresponding to 25 kt, 30 kt, or 35 kt winds are surpassed. Note the same correction throughout the forecast period as in Table 4 presumes the model-predicted tendencies are correct and any forecast timing error arises from an incorrect initial value.

The LTVs of relative vorticity and warm core for the NOGAPS model corresponding to initial CARQ intensities of 15 kt to 35 kt are indicated in Figure 3 and Figure 4, respectively. In the first post-processing test of the impact of applying these LTVs as the initial values, the NOGAPS forecasts of 850 - 500 mb relative vorticity and 700 - 500 mb warm core were examined for the times at which the LTV values corresponding to 35 kt were exceeded. This 35 kt threshold was selected for the first test because it will have the largest number of forecasts since all forecasts with initial CARQ intensity values from 15 kt to 30 kt will be included.

Also indicated in Figure 1, the relative vorticity (or warm core value) in the initial NOGAPS analysis may be either larger or smaller than the LTV corresponding to the CARQ intensity. Given that these NOGAPS forecasts are drawn from the same development sample as was used to derive the LTVs, approximately one-third of the initial NOGAPS analyses are expected to have values less than the LTV. In the post-processing application to the NOGAPS summarized in Table 5, 26 (25) of the initial 850 - 500 mb relative vorticities (700 - 500 mb warm cores) were smaller than the LTV, and 61 (62) were greater than the LTV, so the one-third and two-thirds distribution is only approximate.

Table 5. Distributions as a function of initial CARQ intensity (kt) of the initial 850 - 500 mb relative vorticity (ζ_{init}) and 700 - 500 mb warm core (WC_{init}) from the VORTRACK files of NOGAPS forecasts in three categories: $\zeta_{init} > \zeta_{LTV} < \zeta_{35}$ (downward ζ adjustment), $\zeta_{init} > \zeta_{LTV} > \zeta_{35}$ (false alarm), and $\zeta_{init} < \zeta_{LTV}$ (upward ζ adjustment) in left columns and $WC_{init} > WC_{LTV} < WC_{35}$ (downward adjustment), $WC_{init} > WC_{LTV} > WC_{35}$ (false alarm), and $WC_{init} < WC_{LTV}$ (upward adjustment) in right columns. Whereas an upward or downward adjustment in the initial ζ or WC during the post-processing step in Figure 1 will potentially lead to an improved forecast of the F(TC) threshold-passing time, the initial values in the false alarm cases already exceed the 35 kt LTV thresholds. The threshold LTV values for ζ in 10^{-5} s^{-1} units (WC in °C) for NOGAPS at 15, 20, 25, 30, and 35 kt are: 1.61 (0.05), 2.32 (0.10), 3.03 (0.15), 3.74 (0.20), and 4.45 (0.25), respectively.

Initial CARQ Intensity	850 - 500 mb ζ			700 - 500 mb warm core		
	$\zeta_{init} > \zeta_{LTV}$		$\zeta_{init} < \zeta_{LTV}$	$WC_{init} > WC_{LTV}$		$WC_{init} < WC_{LTV}$
	$\zeta_{init} > \zeta_{LTV} < \zeta_{35}$	$\zeta_{init} > \zeta_{LTV} > \zeta_{35}$		$WC_{init} > WC_{LTV} < WC_{35}$	$WC_{init} > WC_{LTV} > WC_{35}$	
15	0	0	0	0	0	0
20	21	0	6	10	9	8
25	18	5	11	10	13	11
30	2	15	9	2	18	6
Total	41	20	26	22	40	25

For those initial NOGAPS 850 - 500 mb relative vorticities or 700 - 500 mb warm core values in Table 5 that are greater than the LTVs, some values actually exceed the LTV threshold value for a 35 kt tropical storm, which may be considered as a false alarm since the CARQ intensity for all these cases is 30 kt or less. Perhaps the most surprising result of the post-processing of the NOGAPS forecast is the large fraction of false alarms, which are in the $\zeta_{init} > \zeta_{LTV} > \zeta_{35}$ and $WC_{init} > WC_{LTV} > WC_{35}$ columns in Table 5. Notice that 40 (20) of the 97 NOGAPS initial analyses have false alarms for the 700 - 500 mb warm core (850 - 500 mb relative vorticities). As expected, many of these false alarms occur for initial CARQ intensities of 30 kt since these storms are close to the threshold. However, nine of the 40 warm core false alarms occur with a CARQ intensity of 20 kt, which means a large negative correction must be applied to the initial 700 - 500 mb warm core values in these NOGAPS forecasts.

One might have expected that those initial NOGAPS values of relative vorticities and warm cores that were too large (small) would have been grouped at the higher (lower) CARQ values. However, half of those $\zeta_{init} > \zeta_{LTV}$ cases that are not false alarms occur for the CARQ intensity of 20 kt (Table 5), which suggests the NOGAPS initial relative vorticities are typically high for these weaker disturbances. By contrast, many of the cases that require an upward adjustment in relative vorticity fall in the 25 kt (11 of 26 cases) and 30 kt (9 of 26 cases) categories. The non-false alarm warm core cases requiring adjustments upward or downward tend to fall almost equally in the CARQ intensities of 20 kt and 25 kt categories (Table 5).

A schematic of the various forecast timing errors that may arise from model forecasts with unadjusted and adjusted initial conditions as described above is given in Figure 8. If the initial 850 - 500 mb relative vorticity in the unadjusted NOGAPS forecast exceeds the threshold LTV for 35 kt, then the remainder of that forecast must have a negative relative vorticity correction in hopes that an Adjusted Hit occurs. Similarly, if an unadjusted NOGAPS forecast is a Miss because it started from a too small relative vorticity and the predicted 850 - 500 mb relative vorticity never exceeded the threshold LTV for 35 kt, then a positive relative vorticity correction must be applied.

Whereas such an initial adjustment may then lead to an Adjusted Hit, it is also possible that the adjustment is not large enough for the forecast to exceed the 35 kt threshold LTV, so this is labeled as an Adjusted Miss in Figure 8. Other post-processing technique adjustments to the NOGAPS initial 850 - 500 mb relative vorticity to agree with the CARQ intensity will shift the forecast time at which the threshold LTV for 35 kt is exceeded.

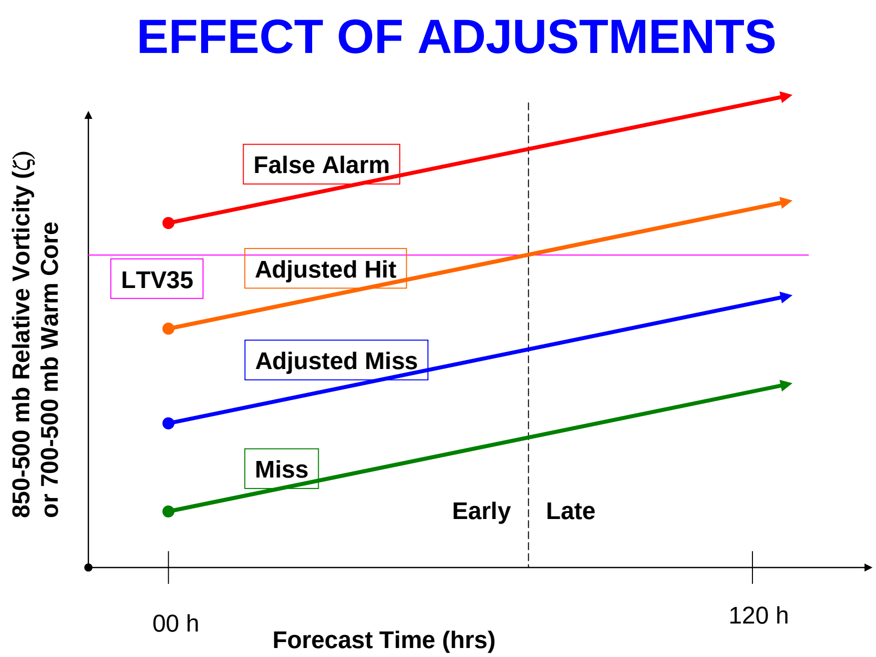


Figure 8 Schematic of the forecast evaluations for unadjusted and adjusted initial conditions that result in false alarms (exceeds LTV_{35} from the beginning) or misses (does not exceed LTV_{35} during the forecast interval) versus adjusted initial conditions that may result in a hit or a miss. For both the unadjusted and the adjusted forecasts, the forecast time may be either early or late relative to the verification time (CARQ intensity equal to 35 kt).

Therefore, the unadjusted versus adjusted forecast evaluations in Tables 6, 7, 9, 10, 12, and 13 will also include Hit categories of early, on time, or late relative to the verification time at which the Best Track intensity was equal to 35 kt. These forecast timing differences due to the adjusted initial condition may lead to a Better (B), Tie (T), or Worse (W) forecast. For example, the unadjusted NOGAPS warm core forecast in Table 4 predicts the $LTV_{35} = 0.30$ would be exceeded by 12 hours and the adjusted

NOGAPS forecast would exceed the LTV_{35} threshold after only 6 h. Since the Best Track intensity = 35 kt occurred at Hour 36, the adjusted forecast is actually worse by 6 h. It is also possible that an unadjusted forecast that was a Hit is degraded by the initial condition adjustment such that it becomes a Miss.

The most significant conclusion from the evaluation of the NOGAPS unadjusted and adjusted forecasts of 850 - 500 mb relative vorticity in Table 6 is that 52 unadjusted forecasts that were Misses (did not exceed the 35 kt LTV threshold) continued to be misses after the post-processing technique in Figure 1 was applied. An example of this unadjusted Miss forecast of 850 - 500 mb relative vorticity continuing as an adjusted Miss forecast is given in Table 4. Clearly an initial relative vorticity adjustment can not overcome an incorrect negative model forecast vorticity tendency. The 52 cases in the unadjusted Miss / adjusted Miss category of Table 6 indicate this is a common problem for NOGAPS forecasts of 850 - 500 mb relative vorticity. Another significant conclusion from Table 6 is that 18 of the 20 false alarms in terms of 850 - 500 mb relative vorticity (Table 5) became Misses after the post-processing adjustment. Two Hit-early and two Hit-late unadjusted NOGAPS forecasts were also degraded to become Misses. Only 11 of the 87 NOGAPS forecasts became a Better or were a Tie with the unadjusted forecasts, and six adjusted NOGAPS forecasts were degraded relative to the unadjusted forecasts. Even when the post-processing led to a better forecast, these adjusted NOGAPS were generally late relative to verifying Best Track intensity of 35 kt.

By contrast, the post-processing adjustment of the NOGAPS 700 - 500 mb warm core forecasts generally led to improved timing of when the LTV_{35} was predicted relative to the unadjusted NOGAPS forecasts (Table 7). A notable change in the 87 unadjusted NOGAPS warm core forecasts was that only six were Misses compared to the 56 NOGAPS 850 - 500 mb relative vorticity forecasts (Table 6), so NOGAPS does better in the tendency forecast of the warm core. As summarized in Table 5, the initial warm core values in NOGAPS are frequently (40 of 87 forecasts) too high, which are listed as false alarms in Table 7. The post-processing adjustment for 26 of these false alarm cases resulted in an improvement relative to the unadjusted forecasts in predicting the time that the LTV_{35} would be exceeded. Whereas the post-processing of the other 14 warm core

false alarms led to 14 Miss forecasts, it is unlikely that this is due to an improper adjustment in the initial conditions, and is more likely or due to model prediction error.

Table 6. Summary of the NOGAPS 850 - 500 mb relative vorticity forecast errors for the timing of the CARQ intensity of 35 kt is starting from various categories of unadjusted initial conditions versus the post-processing adjustment of the initial conditions and an indication is given as to the numbers of adjusted forecasts that were Better (B), a Tie (T), or Worse (W) than the unadjusted forecasts. The definitions of the terms in the forecast evaluation are indicated in the schematic in Figure 8, and an indication is given as to the numbers of adjusted forecast that were Better (B), a Tie (T), or Worse (W) than the unadjusted forecasts.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms			B (2)	18
Hit-early				W (2)
Hit-on time				
Hit-late		B (1)	B (2); T (2); W (2)	W (2)
Misses	B (1)		B (3)	52

Table 7. Summary of NOGAPS forecast timing errors for a CARQ intensity of 35 kt as in Table 6, except for the 700 - 500 mb warm core threshold.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms	B (18)	B (2)	B (6)	14
Hit-early	B (3); T (13); W (5)	B (2)		W (1)
Hit-on time	W (2)	T (3)	W (1)	
Hit-late	B (3); W (4)	B (1)	B (1); T (2)	
Misses	B (2)			4

For the 41 unadjusted NOGAPS warm core forecasts that were classified as Hits in Table 7, 10 were better, 18 were ties, and 13 were worse (counting one Miss) of the adjusted NOGAPS forecasts by at least 6 h in the timing of when the LTV_{35} would be

exceeded. One example of an adjusted NOGAPS warm core forecast that was 6 h worse than the unadjusted NOGAPS forecast is given in Table 4. Taking into account the large number of false alarm cases that became Hits after the post-processing adjustment of the initial warm core values, the post-processing technique was successful for the NOGAPS warm core forecasts.

B. GFS FORECASTS AFTER POST-PROCESSING

The distributions of the initial 850 - 500 mb relative vorticity and the 700 - 500 mb warm core values for the GFS model forecasts as a function of the initial CARQ intensities are given in Table 8. The LTV Final relative vorticities and warm core values for the CARQ intensities from 15 kt to 35 kt are given in Figure 3 and Figure 4, respectively. As in the NOGAPS test above, the post-processing test of the impact of applying these LTV values as an adjustment to the initial GFS relative vorticity and warm core will only be for the timing forecast errors of when the LTV ζ_{35} and WC_{35} values will be exceeded.

As in the NOGAPS distribution (Table 5), the GFS distribution in Table 8 indicates there was a surprisingly large number of the initial warm core (34 cases in the $WC_{init} > WC_{LTV} > WC_{35}$ column) and relative vorticity (21 cases in the $\zeta_{init} > \zeta_{LTV} > \zeta_{35}$ column) values for the GFS model are false alarms. About 40% (55%) of the warm core (relative vorticity) false alarms occur when the CARQ intensity is 30 kt, so these values may only slightly exceed the corresponding LTV_{30} values.

The GFS model was integrated for four cases when the CARQ intensity was only 15 kt (Table 8), and in all four cases the initial relative vorticity exceeded the LTV_{15} value in Figure 4. Excluding the false alarms, the initial relative vorticities and warm core values that exceeded (i.e., requiring a downward adjustment) or were below (i.e., requiring an upward adjustment) the 20 kt and 25 kt CARQ intensity LTVs were roughly balanced.

In common with NOGAPS, 21 of the GFS 850 - 500 mb relative vorticity forecasts began with initial conditions that were categorized as false alarms (Table 8). Applying the post-processing adjustment to the initial conditions of these GFS forecasts resulted in 10 Hits and 11 Misses (Table 9), which is a large improvement over the NOGAPS post-processing in which 18 of the 20 were Misses. The unadjusted GFS relative vorticity forecasts included 32 Misses (compared to 56 Misses for NOGAPS),

Table 8. As in Table 5, except for distributions as a function of initial CARQ intensity (kt) of the initial 850 - 500 mb relative vorticity (ζ_{init}) and 700 - 500 mb warm core (WC_{init}) from the VORTRACK files for GFS forecasts in three categories. The threshold LTV values for ζ in 10^{-5} s^{-1} units (WC in °C) for GFS at 15, 20, 25, 30, and 35 kt are: 2.69 (0.06), 3.15 (0.12), 3.61 (0.18), 4.07 (0.25), and 4.53 (0.33), respectively.

Initial CARQ Intensity	850 - 500 mb ζ			700 - 500 mb warm core		
	$\zeta_{init} > \zeta_{LTV}$		$\zeta_{init} < \zeta_{LTV}$	$WC_{init} > WC_{LTV}$		$WC_{init} < WC_{LTV}$
	$\zeta_{init} > \zeta_{LTV} < \zeta_{35}$	$\zeta_{init} > \zeta_{LTV} > \zeta_{35}$		$WC_{init} > WC_{LTV} < WC_{35}$	$WC_{init} > WC_{LTV} > WC_{35}$	
15	4	0	0	2	0	2
20	11	1	8	12	4	4
25	19	9	16	11	16	16
30	7	11	9	4	14	10
Total	41	21	33	29	34	32

and only five of these were better after the post-processing adjustment of the initial conditions. While this was a better performance than for the NOGAPS post-processing of Misses (only four out of 56), it still indicates that an under-prediction of the 850 - 500 mb relative vorticity tendency by a model is less likely to be improved by the post-processing technique.

Table 9. Summary of the GFS 850 - 500 mb relative vorticity forecast timing errors in a format similar to Table 6.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms	B (2)	B (2)	B (6)	11
Hit-early	B (3); T (4); W (6)	B (1)	W (3)	W (5)
Hit-on time		T (6)		W (1)
Hit-late	W (1)		B (4); T (7); W (1)	
Misses	B (4)		B (1)	27

Table 10. Summary of the GFS 700 - 500 mb warm core forecast timing errors in a format similar to Table 7.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms	B (15)	B(3)	B (4)	13
Hit-early	B (7); T (13); W (9)		W (1)	
Hit-on time	W (1)	T (1)	W (2)	
Hit-late	B (2); T (1)	B (4)	T (8); W (4)	
Misses	B (1)			6

A somewhat mixed performance of better and worse timing errors was achieved by post-processing those unadjusted GFS relative vorticity forecasts that were Hits. For example, the Hit-early forecasts resulted in four with better and four with ties, but 14 had worse timing of the LTV_{35} threshold being exceeded, including five that became Misses. These forecasts that became Misses are often just below the LTV_{35} threshold, which

indicates a problem with using a discrete value (that is not well known) to distinguish between a Hit and a Miss. By contrast, having four better and seven ties with only two worse forecasts after adjustment of the Hit-late forecasts is a positive result.

As was the case with the NOGAPS forecasts of the 700 - 500 mb warm core, a large fraction (34 of 95 cases or 37%) of the GFS warm core forecasts began with initial conditions that were classified as false alarms (Table 10). After the post-processing adjustment, 22 of these GFS forecasts were improved (Table 10). Both the unadjusted GFS and NOGAPS forecasts of the warm core had few Misses (seven and six, respectively) and one forecast by GFS and two forecasts by NOGAPS were improved by the post-processing adjustment. Both the unadjusted GFS and NOGAPS warm core forecasts had a large fraction of Hit-early timing errors, and the post-processing adjustment for both models resulted in a mixed performance. For the GFS adjusted forecasts, seven were better, 13 were ties, and 10 had a worse timing error by at least 6 h of the time that the LTV_{35} warm core threshold would be exceeded. A somewhat mixed performance was also achieved for the unadjusted Hit-late forecasts, with six better, nine ties, and four worse cases after the post-processing adjustment of the initial conditions (Table 10). Taken together with the large number of the GFS warm core forecasts that began from false alarms, the post-processing adjustment results in improved forecasts of the timing of when the LTV_{35} threshold will be exceeded.

C. UKMO FORECASTS AFTER POST-PROCESSING

The distributions of the initial 850 - 500 mb relative vorticity and the 700 - 500 mb warm core values for the UKMO model forecasts as a function of the initial CARQ intensities are given in Table 11. The LTV Final relative vorticities and warm core values for the CARQ intensities from 15 kt to 35 kt are given in Figure 3 and Figure 4, respectively. As in the NOGAPS and the GFS tests above, the post-processing test of the impact of applying these LTVs as an adjustment to the initial UKMO relative vorticity and warm-core values will only be for the timing forecast errors of when the LTV_{35} and WC_{35} values will be exceeded.

Table 11. As in Table 5, except for distributions as a function of initial CARQ intensity (kt) of the initial 850 - 500 mb relative vorticity (ζ_{init}) and 700 - 500 mb warm core (WC_{init}) from the VORTRACK files of UKMO forecasts in three categories. The threshold LTV values for ζ in 10^{-5} s^{-1} units (WC in °C) for UKMO at 15, 20, 25, 30, and 35 kt are: 1.13 (0.01), 2.09 (0.08), 3.17 (0.15), 4.37 (0.22), and 5.53 (0.30), respectively.

Initial CARQ Intensity	**850 - 500 mb ζ**			**700 - 500 mb warm core**		
	$\zeta_{init} > \zeta_{LTV}$		$\zeta_{init} < \zeta_{LTV}$	$WC_{init} > WC_{LTV}$		$WC_{init} < WC_{LTV}$
	$\zeta_{init} > \zeta_{LTV} < \zeta_{35}$	$\zeta_{init} > \zeta_{LTV} > \zeta_{35}$		$WC_{init} > WC_{LTV} < WC_{35}$	$WC_{init} > WC_{LTV} > WC_{35}$	
15	2	0	0	0	0	2
20	16	1	8	10	3	12
25	19	8	13	9	23	8
30	6	9	7	2	15	5
Total	43	18	28	21	41	27

Table 12. Summary of the UKMO 850 - 500 mb relative vorticity forecast timing errors in a format similar to Table 6.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms		B (1)	B (7)	10
Hit-early	B (4); T (1)			W (2)
Hit-on time			W (1)	
Hit-late	W (1)		T (6); W (5)	W (3)
Misses	B (1)	B (2)	B (6)	39

As in the NOGAPS (Table 5) and GFS (Table 8) distributions, a surprisingly large number of the initial warm core (41 cases in the $WC_{init} > WC_{LTV} > WC_{35}$ column) and relative vorticity (18 cases in the $\zeta_{init} > \zeta_{LTV} > \zeta_{35}$ column) values for the UKMO model are false alarms. Whereas 50% (about 37%) of the relative vorticity (warm core) false alarms occur when the CARQ intensity is 30 kt, the remaining false alarms occur for intensities of 20 kt (especially more than half of the warm core) and 25 kt. Excluding the false alarms, the initial relative vorticities and warm core values that exceeded (i.e., required a downward adjustment) or were below (i.e., requiring an upward adjustment) the 20 kt and 25 kt CARQ intensity LTVs were roughly balanced.

The most significant result for the UKMO 850 - 500 mb relative vorticity was a large number (48 of 89) of unadjusted forecasts that were Misses, and relatively few (nine) were improved by the post-processing technique (Table 12). This performance with missed relative vorticity forecasts is similar to the GFS (only five of 32 improved) and somewhat better than for the NOGAPS (only four of 56 improved). Only eight of the 18 UKMO relative vorticity forecasts that began with initial conditions that were categorized as false alarms were improved by the post-processing technique (Table 12). The remaining 10 UKMO relative vorticity forecasts that were false alarms became Misses, which is a similar performance for the GFS (11 cases) and the NOGAPS (18 cases). The post-processing technique was not particularly successful for the unadjusted UKMO relative vorticity forecasts that were Hits. After the post-processing adjustment, only four were better timing forecasts and seven were Ties, but 12 were made worse including five that became Misses. This poor UKMO performance was similar to the NOGAPS relative vorticity forecasts that were Hits (three better, two ties, and six made worse), whereas the adjusted GFS forecasts had eight better, 17 ties, and 17 made worse.

Just as strongly as the NOGAPS (40 of 87 forecasts or 46%) or the GFS (34 of 95 or 37%) forecasts, a large fraction (41 of 89 or 46%) of the UKMO 700 - 500 mb warm core forecasts started from initial conditions classified as false alarms (Table 13). Similar to the NOGAPS and GFS models, the post-processing adjustments of the UKMO warm core forecasts led to a large number (23) of Hits (Table 13). The unadjusted UKMO warm core forecasts had proportionally more (16 of 89 or 18%) Misses than either the

NOGAPS or GFS forecasts, and only three of these were converted to Hits after the post-processing adjustment. The performance of the post-processing for those unadjusted UKMO warm core forecasts classified as Hits was again mixed (eight better, 14 ties, and ten made worse), as was the case for the NOGAPS and GFS models. However, including the 23 false alarm cases that became Hits, the post-processing leads to a positive impact in forecast timing of when the warm core LTV_{35} is exceeded.

Table 13. Summary of the UKMO 700 - 500 mb warm core forecast timing errors in a format similar to Table 7.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms	B (13)		B (10)	18
Hit-early	B (5); T (11); W (2)		B(1); T (1)	W (2)
Hit-on time	W (2)	T (1)	W (1)	
Hit-late	W (2)	B (1)	B (1); T (1); W (1)	
Misses	B (2)	B (1)		13

D. SUMMARY OF NOGAPS, GFS, AND UKMO FORECASTS AFTER POST-PROCESSING

A general conclusion of the post-processing technique applied to the NOGAPS, GFS, and UKMO model predictions of the 850 - 500 mb relative vorticity is that it is only marginally successful. While some of the false alarms are improved, the lack of success for the unadjusted forecast Misses (particularly for the NOGAPS and somewhat for UKMO), and that many of the unadjusted Hits were shifted to Late or Misses, suggests that the model-predicted relative vorticity tendencies were generally too small. If the model-predicted tendencies were unbiased, the post-processing technique for the false alarms and for the Hits would have been expected to have a more balanced distribution of early and late improvements rather than be biased toward Late forecast timing, or even Misses. Similarly, the large fraction of the unadjusted model forecasts that are Misses,

and that then adjusting the initial relative vorticities upward did not result in many improvements, are further indications that the model-predicted relative vorticity tendencies are biased too small.

The general conclusion of the post-processing technique applied to the NOGAPS, GFS, and UKMO model predictions of the 700 - 500 mb warm core is more favorable. Between 60-65% of the unadjusted model predictions that began as false alarms were converted to Hits in the sense that the predicted warm core LTV_{35} threshold was exceeded during the 120 h forecast interval. Those unadjusted warm core forecasts that were classified Hits were found after the post-processing adjustment to have many cases of the same timing (i.e., were ties) and slightly fewer better forecasts than worse forecasts by at least six hours. The fraction of unadjusted warm core forecasts that were Misses was much smaller than for the relative vorticity forecasts, and the post-processing adjustment of the initial condition only converted a small fraction of these Misses to become Hits. Whereas a low bias in the model-predicted tendencies was concluded to be a significant contributor in the poor performance of the post-processing technique for the 850 - 500 mb relative vorticity, it is not clear that a marked too-low bias exists for the 700 - 500 mb warm core.

E. CONVERTING THE ADJUSTED MODEL FORECAST RELATIVE VORTICITY OR WARM CORE TO A WIND SPEED FORECAST

In the previous post-processing technique approach, the objective was to adjust the initial condition to correspond to the LTV for the CARQ intensity and then to calculate the time at which either the adjusted forecast of 850-500mb relative vorticity or 700 - 500 mb warm core exceeded a LTV threshold corresponding to a CARQ intensity of 35 kt (Figure 9, left side). In examining each adjusted forecast, it became evident that many of the forecast values approached the threshold value, but did not exceed it, and thus they were classified as Misses. This tendency was especially evident when the unadjusted forecast was a Miss and the post-processing adjustment of the initial conditions was insufficient for the forecast value to exceed the threshold. However, several other adjusted forecasts that became Misses were found to be due to the forecast

values now just missing the threshold. Furthermore, the timing of the forecast Hit was often markedly changed when the unadjusted forecast value was quite close to, or just exceeded the threshold value, so that only a small adjustment could either cause the threshold value to be exceeded, or not be exceeded, and thus change the forecast at the time to be classified as a Hit, or not a Hit. This sensitivity is a direct result of setting a threshold value (which may not be known that precisely) that leads to a yes/no advice (e.g., in this case that a value corresponding to 35 kt was exceeded, which may be regarded as a "formation").

Useful information to the forecaster is contained in those cases in which the adjusted model forecast almost exceeded the 35 kt (or 25 kt or 30 kt) threshold. Thus, an additional procedure is proposed to display the maximum wind speeds corresponding to the adjusted model forecast values of 850 - 500 mb relative vorticity and of 700 - 500 mb warm core. Thus, the complete forecast evolution(s) of the wind speed will be displayed, rather than providing the forecaster with a "Hit forecast" that a threshold value was exceeded at time t in the forecast. By displaying the three adjusted model forecasts of wind speed, the forecaster can assess the variability among the solutions and better assess the risk, and after becoming familiar with the model performance in various scenarios will be able to add value to the model predictions. Another useful display would be the consensus, and the spread about the consensus mean, of the maximum wind speed evolution, because the accuracy of a consensus of three skillful models is typically greater than the accuracy of the individual models. Thus, the consensus is often good guidance for the forecaster.

The proposed technique (Figure 9, right side) for forecasting the maximum wind speed each 6 h is to convert the six-hourly adjusted model forecast values of 850 - 500 mb relative vorticity and of 700 - 500 mb warm core by interpolating the maximum wind speeds (CARQ intensities along the abscissa) in Figure 3 and Figure 4, respectively. For example, if the adjusted NOGAPS model forecast relative vorticity after say t hours happened to be equal to $3.88 \times 10^{-5}\ s^{-1}$, the corresponding maximum wind speed (intensity) from Figure 3 would be 32 kt. Since the curves in Figure 3 and Figure 4 are essentially linear, interpolation to intermediate wind speeds will be easily accomplished. For an immediate test, these linear relationships will need to be extrapolated to lower and higher maximum wind speeds (at least 10-40 kt) to allow for larger adjusted model relative vorticities or warm cores that will correspond to maximum wind speeds beyond 35 kt.

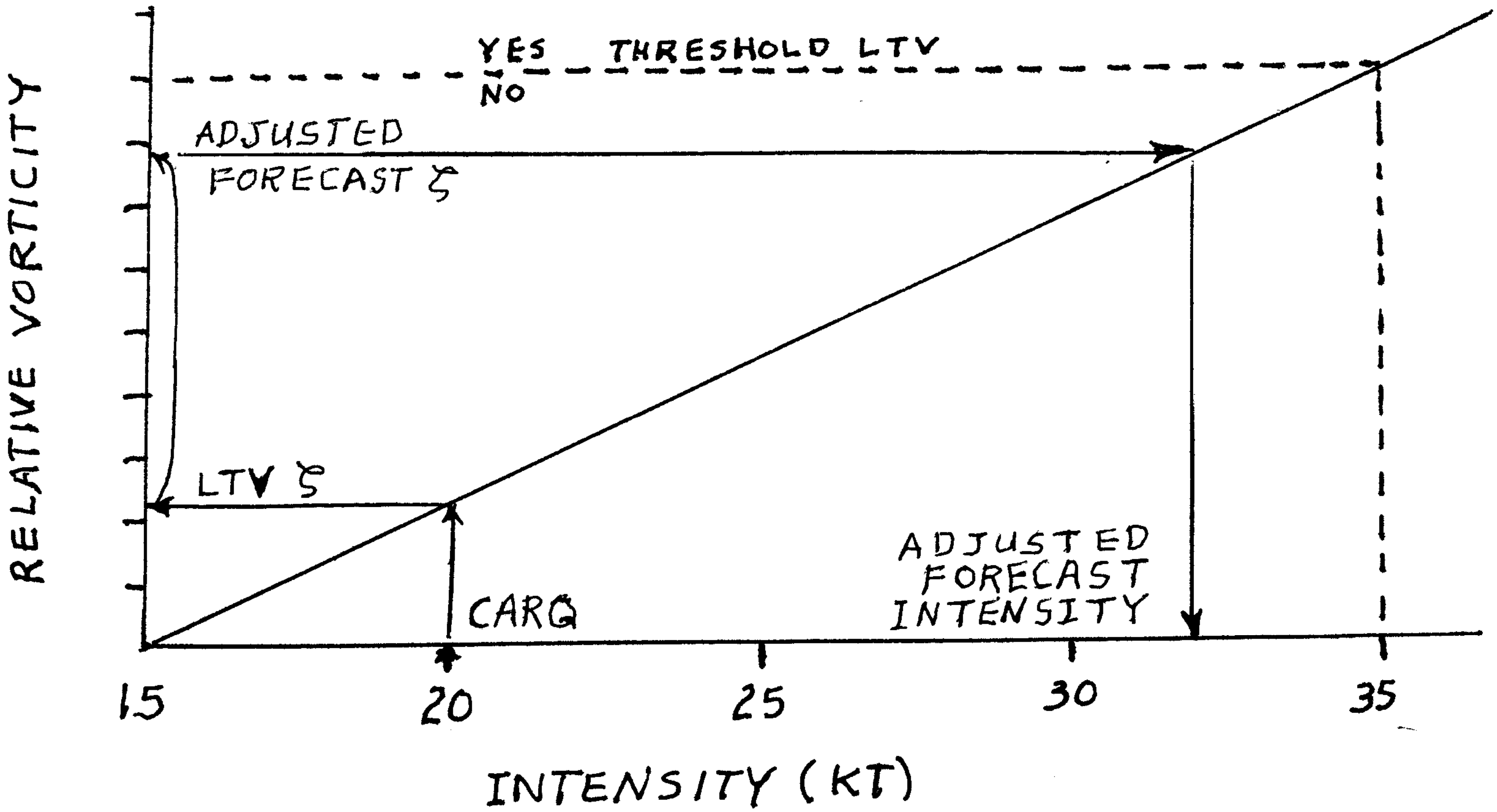


Figure 9 Schematic of the conversion of an adjusted model forecast of 850 - 500 mb relative vorticity (ζ) along the ordinate to an adjusted forecast intensity along the abscissa using the relationship between CARQ intensities and the model LTV relative vorticities as in Figure 3. In the scenario, the CARQ intensity equal to 20 kt is used to specify the initial relative vorticity for the adjusted forecast relative vorticity, which does not exceed the LTV_{35} threshold, and the adjusted forecast intensity is 32 kt.

Since separate maximum wind speeds will be available from the adjusted model forecasts of 850 - 500 mb relative vorticity and of 700 - 500 mb warm core for the three models, six intensity estimates will be available. This situation is a natural for calculating a consensus mean forecast and a spread about that mean. Given knowledge of the relative accuracy of intensity estimates from the two variables in the three models, a weighted consensus mean could be calculated in which the model variable with the greatest skill is given a larger weight (e.g., inversely proportional to the root-mean square intensity).

One might inquire why the forecaster does not just examine the various model outputs for the maximum surface wind speed and use that speed for their warnings. While the forecasters do examine model output, as stated previously tropical cyclones form over oceans where surface wind observations are scarce, except perhaps when a scatterometer pass is available. By contrast, most estimates of maximum wind speed in a tropical disturbance are based on satellite imagery interpretation, which then becomes the basis for the CARQ intensity. As demonstrated above, the model estimates of the wind/vorticity structure of warm core in the tropical disturbance or depression may depart significantly from the CARQ estimate. If one then accepts the satellite-based CARQ estimate as the best intelligence (and most consistent in time) as to the maximum wind speed, the proposed procedure is an indirect method to proceed from the CARQ intensity as an initial value through the model-predicted tendencies of relative vorticity or warm core to a series of values each 6 h that are converted back to forecast intensities. One advantage of this method of deriving the intensity from the model output is that it is completely objective so the forecaster can quickly assess the model guidance.

Using a program developed by Mark Boothe, the intensities corresponding to the unadjusted and adjusted 850 - 500 mb relative vorticity forecasts and 700 - 500 mb warm core forecasts by each of the three global models (NOGAPS, GFS, and UKMO) were calculated. When at least two of the model forecasts were available, a consensus of the intensities (at least four and as many as six) at each forecast time was calculated.

An example of the consensus of unadjusted forecasts when only the NOGAPS and UKMO models were available is given in Figure 10a. The converted intensity values corresponding to the unadjusted relative vorticity forecasts by the two models are somewhat similar, with the UKMO model predicting at a slow increase to 30 kt by 60 h and the NOGAPS model predicting a slow decrease to about 20 kt. If the consensus contained only these two unadjusted relative vorticity values, the result would be a constant intensity, whereas Tropical Depression Arlene actually intensified to 35 kt at time t + 30 h, and further increased to 50 kt by t + 60 h. By contrast, the converted intensity values corresponding to the unadjusted 700 - 500 mb warm core forecasts by the NOGAPS (UKMO) model exceeded (nearly exceeded) the 35 kt threshold already by t + 12 h. Notice that the unadjusted NOGAPS-predicted intensities vary rapidly from 65 kt after only t + 12 h to back to less than 10 kt by t + 24 h, and then to about 50 kt by t + 36 h. Clearly those converted intensities from Figure 4 are based on extrapolated values well beyond the original CARQ intensities. Whereas the unadjusted UKMO-predicted intensities from the warm core forecast vary more smoothly, the converted intensity values are clearly too large after about t + 12 h. If the rapidly varying NOGAPS-predicted intensities from the warm core were not included in the unadjusted consensus, the intensity would have been forecast to hit the 35 kt threshold at t + 36 h versus the actual t + 30 h (Figure 10a).

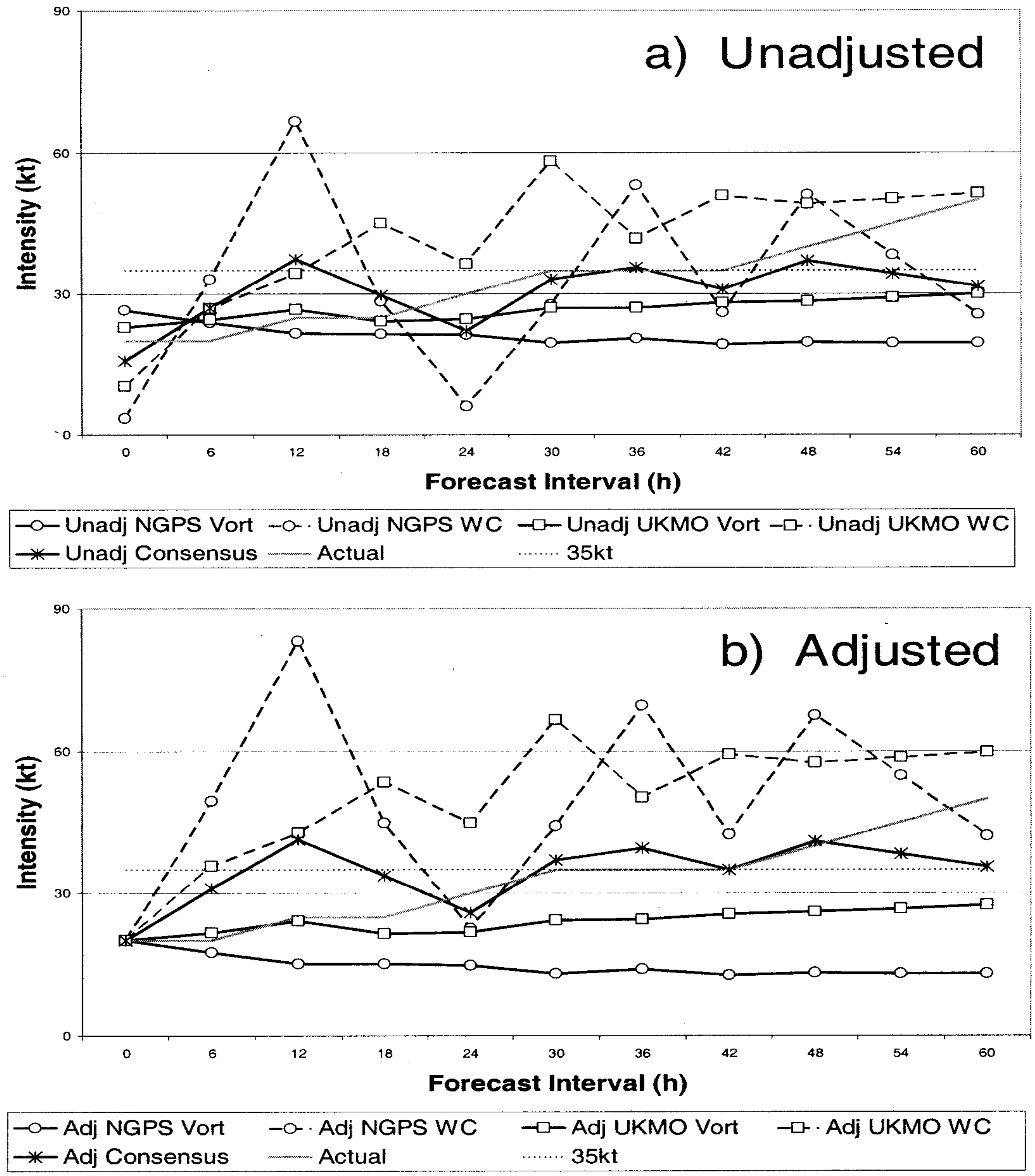


Figure 10 Intensity forecasts converted from the (a) unadjusted and (b) adjusted NOGAPS 850 - 500 mb relative vorticity and 700 - 500 mb warm core forecasts (see inset) from 0000 UTC 8 June 2005 for Tropical Depression Arlene. The consensus of these intensity forecasts and the verifying CARQ intensities are also indicated.

The corresponding consensus of adjusted NOGAPS and UKMO model forecasts is given in Figure 10b. Notice that all of these forecasts begin at the CARQ intensity of 20 kt. This initial condition adjustment led to only small changes in the intensities converted from the relative vorticity forecasts. However, the excessive intensity at 12 h from the adjusted NOGAPS-predicted warm core is made worse, so the 35 kt threshold is forecast to be exceeded by t + 3 h. Likewise, the initial condition adjustment of the UKMO-predicted intensity also made all of the intensities slightly larger. By coincidence, the adjusted consensus intensity forecast surpasses 35 kt at almost the exact time (t + 30 h in Figure 10b) if the excessive consensus intensity at 12 h that is caused by the NOGAPS warm core forecast is ignored. The capability to display the complete forecast intensity evolution for the individual models and both predicted variables along with the consensus intensity forecast can give the forecaster insight about the variability, and provides more information than a yes/no that a LTV threshold value or the converted 35 kt threshold has been exceeded.

For comparison of the consensus approach with the earlier evaluations of the models for predicting the timing that a LTV_{35} threshold is exceeded (Tables 6, 7, 9, 10, 12, and 13), the consensus intensity forecasts of the time that 35 kt is exceeded are evaluated (Table 14). Of the 33 unadjusted forecasts that were false alarms (initial intensity exceeded 35 kt), 20 became Hits and the downward adjustment of the initial intensity also resulted in 13 Misses. As in the cases of the previous model evaluations, few (in this case only two) of the 20 unadjusted consensus forecasts that were Misses (never achieved 35 kt) were improved by the initial intensity adjustment. Likewise, a mixed performance was achieved by the initial intensity adjustment of the consensus of unadjusted forecasts that were Hits, only seven were better, 21 were ties, and 13 were made worse (including six that become Misses) by at least 6 h. When the 20 false alarms that became Hits are included, it is concluded that the post-processing technique of initial intensity adjustment does improve the consensus intensity forecasts of the time at which the 35 kt threshold is exceeded. However, it is again evident the uncertainty in the initial intensity value is not the only limitation in forecasting the time that an intensity of 35 kt

will be exceeded. The model forecasts of the relative vorticity and (especially for NOGAPS) the warm core also contributes to these timing forecast errors.

Table 14. Summary of the unadjusted consensus versus the adjusted consensus forecast errors for the timing of when the CARQ intensity of 35 kt is exceeded with an indication as to the numbers of adjusted forecasts that were Better (B), a Tie (T), or Worse (W) than the unadjusted forecasts in a format similar to Table 6.

Unadjusted Initial Conditions	**Adjusted Initial Conditions**			
	HITS			MISSES
	Early	On-time	Late	
False alarms	B (6)	B (3)	B (11)	13
Hit-early	B (4); T (10); W (2)	B (1)	T (1)	W (5)
Hit-on time	W (2)	T (2)		
Hit-late	B (1)	B (1)	T (8); W (3)	W (1)
Misses		B (1)	B (1)	18

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSIONS AND RECOMMENDATIONS

Since 2003 when the NHC and JTWC started issuing tropical cyclone track forecasts through 120 h, an urgent requirement has been to improve formation and intensification forecast skill because a tropical cyclone can form and intensify to a hurricane within five days. While some conjecture exists as to whether the global models can provide forecasters guidance regarding tropical cyclone formation, the global models are usually the only source during the early stages when regional models are not integrated. Due to the lack of conventional observations over tropical oceans, satellite imagery interpretation is the major tool for analyzing and short-term forecasting of tropical cyclone formation. Therefore, the goal of this study was to test the impact on tropical cyclone formation forecasts with an improved specification of the initial conditions for the three global models and their consensus from using the NHC / TPC CARQ primarily satellite-derived intensities as 'truth' via an adjustment of these initial conditions in a post-processing technique.

A. CONCLUSIONS

1. Numerical Model Approach to TC Formation

Global model forecasts are generally used to predict the synoptic-scale conditions. These numerical models take an initial value approach as in Equation (1) to generate a non-linear solution throughout the forecast period. Model predictions of TC formation are defined here using thresholds of the lower tropospheric (850 – 500 mb layer averaged) relative vorticity and 700 - 500 mb warm core. These model-derived variables have three aspects of uncertainty associated with them.

a. Three Aspects of Uncertainty

Since these global models require accurate initial conditions to predict TC formation, small errors in model initial conditions may lead to large variations in the predictions for formation time. Second, inaccuracies in the model-predicted time

tendencies grow throughout the forecast period. Third, the thresholds in the model-predicted variables that may be associated with a tropical cyclone formation are not well-defined. In this thesis, the objective was to minimize these uncertainties by testing a post-processing of the global model forecasts to improve and extend the time interval for accurate formation predictions. In addition, understanding these sources of uncertainties in global model predictions of tropical cyclone formation will assist in the efforts to mitigate them.

b. Definition of Formation

Cowan (2006) and several other studies have defined TC formation as when the tropical disturbance first appears in Best Track from NHC. Weaknesses in this practice due to variations in intensity upon first appearance are indicated in Figure 2. Here Lower Tercile Values (LTVs) of either relative vorticity or warm core that correspond to the more accepted "official" formation thresholds of 25 kt for tropical depressions or 34 kt for tropical storms are defined from the 2005 Atlantic season.

2. CARQ and VORTRACK Datasets

The basic hypothesis tested here is that CARQ intensity data will significantly improve the tropical cyclone formation forecasts by providing more accurate initial conditions for the global models. Complete synthetic tropical cyclone observations (i.e., bogus vortices) are generally not inserted into the NOGAPS and UKMO global models prior to 35 kt intensities. Therefore, vortex intensities in the global model analyses and forecasts may depart significantly for the CARQ intensities that are satellite-derived using Dvorak techniques with T-numbers less than 2.5. Thus, an improved forecast is expected from the reduction of initialization errors by a post-processing adjustment to make the initial conditions consistent with the CARQ intensity. The objective of this thesis was to document the improvement in the forecasts due to adjustment of the initial conditions for each of the three operational models and also a consensus of the three models.

3. LTVs from CARQ-based Intensities

In this feasibility study, only the same two parameters (850 – 500 mb relative vorticity and 700 – 500 mb warm core) that Cowan (2006) used have been utilized to identify differences between vortices that are forecast to intensify into a tropical cyclone (Hits) and those that do not (Misses).

a. LTVs as Definition of Formation

Both relative vorticity and warm core parameters were analyzed to find consistent LTVs that would correspond to the CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt. Separate LTVs for each of the three models were used as threshold values for determining formation. The second purpose for defining the LTVs was their use in the post-processing technique to adjust the initial values of each model-predicted parameter to correspond with each of the CARQ intensity categories. The 700 - 500 mb warm core values varied more smoothly than the 850 - 500 mb relative vorticity with increasing CARQ intensity as the histograms trended more toward a Gaussian distribution. Both statistical and case methods of estimating LTVs from histograms of the relative vorticity and warm core analyzed values were found to be plausible. For use in the post-processing technique, the preliminary values were smoothed to become the quasi-linear LTV Final values (Figure 3 and Figure 4) with increasing intensity. Variations in the slopes of the LTV Final values from each of the three global models are probably related to the differences in data assimilation, horizontal grid spacing, model physics, parameterization schemes, and other factors. The UKMO model consistently had the largest variation with CARQ intensity, but the differences were smaller for 700 – 500 mb warm core than for 850 – 500 mb relative vorticity.

b. Use of LTVs as 'Truth'

Although either of the LTVs corresponding to 25 kt or 35 kt could be used as the thresholds to define formation time in the models, in this feasibility test only the intensity level of 35 kt is used to see if the models have significant skill in forecasting the transition from an incipient tropical seedling to a tropical depression into a tropical storm.

Significantly smaller numbers of cases with CARQ values less than 25 kt are available in this sample. The test for each model was its skill in forecasting the time that the 35 kt threshold of either relative vorticity or warm core was exceeded. These two 35 kt LTV Final thresholds as given in Figure 3 and Figure 4, and thus are the 'truth' values for evaluating the model-predicted relative vorticity and warm core for each model.

4. Post-processing Technique

In the post-processing technique (Figure 1), the initial 850 – 500 mb relative vorticity or 700 – 500 mb warm core from VORTRACK for each model was adjusted so that the initial value agrees with the LTV for that model and for that particular CARQ intensity value. These adjustments to each model's initial condition were upwards (downwards) if the model initial relative vorticity or warm core was too low (too high) relative to the LTV corresponding to the CARQ intensity given in the curves in Figure 3 and Figure 4. The primary test was whether model-predicted tendencies of 850 – 500 mb relative vorticity and 700 – 500 mb warm core beginning from the adjusted initial value produced more accurate forecasts for when the formation time was exceeded than did the unadjusted model forecast. Each model's forecast timing error was calculated by the number of hours the adjusted initial condition forecast was closer to the actual time the initial 35 kt CARQ intensity occurred when compared to the unadjusted initial condition forecast.

The main conclusion from post-processing of NOGAPS model predictions of 850 - 500 mb relative vorticity was that it was only marginally successful. Some False Alarms were improved upon, but there was significant lack of success for the unadjusted initial condition forecasts that were Misses, since adjusting the initial relative vorticities upward did not result in many improvements. Furthermore, the unadjusted initial condition forecast Hits were shifted too much so that the adjusted forecasts were Hit-Late or Miss forecasts. If the model-predicted tendencies were unbiased, the post-processing technique for the False Alarms and Hits would have been a more balanced distribution of

early and late improvements rather than be biased toward a late forecast timing, or even a Miss. This is a good indication that the NOGAPS model-predicted relative vorticity tendencies are biased too small.

As in the NOGAPS model, a large number (21 cases) of False Alarms occurred in the initial 850 – 500 mb relative vorticity for the GFS model. Most of these False Alarms had initial values that just exceeded the LTV_{35}. After applying the post-processing adjustments to the initial conditions for relative vorticity, 10 adjusted GFS forecasts became Hits, which is a much better performance than for the 20 unadjusted NOGAPS False Alarms that resulted in 18 Misses after the post-processing adjustment. Fewer (32 cases versus 56 cases for NOGAPS) of the unadjusted GFS relative vorticity forecasts were Misses. However, only five of these GFS forecasts were improved after the post-processing of the initial conditions, which again is an indication of under-prediction of the 850 – 500 mb relative vorticity tendency by both GFS and NOGAPS. As was the case with the NOGAPS, mixed results were achieved from the post-processing of the unadjusted GFS forecasts of relative vorticity that were Hits. For example, the Hit-early category had a significant number (14 out of 22) of Worse forecasts of the time that LTV_{35} was exceeded.

As in the NOGAPS (Table 5) and GFS (Table 8) adjusted versus unadjusted relative vorticity forecasts, a large number (18 cases) of the initial values for the UKMO model were False Alarms, and only eight were improved by the post-processing technique (Table 12). Similar to NOGAPS, the most significant result for the UKMO model 850 – 500 mb relative vorticity was a large number (48 cases out of 89 total cases) of unadjusted forecasts that were Misses, and relatively few (nine) were improved by the post-processing technique (Table 12). The post-processing technique was also not particularly successful for the unadjusted UKMO relative vorticity forecasts that were Hits. After the post-processing adjustment, only four forecasts had better timing forecasts, seven were Ties, and 12 were made worse, including five that became Misses. This poor UKMO performance was similar to the NOGAPS relative vorticity forecasts that were Hits (three better, two ties, and six made worse), whereas the adjusted GFS forecasts had eight better, 17 ties, and 17 made worse.

The main conclusion of the post-processing technique applied to the NOGAPS, GFS, and UKMO model predictions of the 700 - 500 mb warm core is more favorable than for the 850 - 500 mb relative vorticity. Between 60-65% of the unadjusted model predictions that began as false alarms were converted to Hits in the sense that the predicted warm core LTV_{35} threshold was exceeded during the 120 h forecast interval. Those unadjusted warm core forecasts that were classified Hits were found after the post-processing adjustment to have many cases of the same timing (i.e., were ties) and slightly fewer better forecasts than worse forecasts by at least six hours. The percentage of unadjusted warm core forecasts that were Misses was much smaller than for the relative vorticity forecasts, and the post-processing adjustment of the initial condition only converted a small fraction of these Misses to become Hits. Whereas a low bias in the model-predicted tendencies was concluded to be a significant contributor in the poor performance of the post-processing technique for the 850 - 500 mb relative vorticity, it is not clear that a marked too-low bias exists for the 700 - 500 mb warm core.

5. Converting Adjusted Model Forecast Back to a Wind Speed Forecast

In the above test of the post-processing technique, the adjustment of the initial condition to correspond to the LTV for the CARQ intensity essentially led to a yes / no decision as to whether the adjusted forecasts of either 850 – 500 mb relative vorticity or 700 – 500 mb warm core exceeded the formation LTV for 35 kt. In many cases, these were close calls in which the threshold was approached but not exceeded. Thus, an additional procedure was used to display the forecasts in terms of the maximum wind speeds that correspond to the unadjusted and the adjusted model forecast values. That is, the entire forecast evolution was displayed in terms of equivalent maximum wind speeds from which the forecaster can easily grasp and assess the variability among the solutions, or use a consensus forecast approach. Separate conversions to equivalent maximum wind speeds are derived for the relative vorticity and warm core forecasts for each model, which yields six intensity estimates. Only a simple consensus mean forecast was tested here, although later a weighted consensus could be calculated from model variables

giving higher weight to those model-predicted variables with greater skill after comparison with the satellite-based interpretation CARQ intensity.

Comparison of the consensus approach with earlier evaluations of models for predicting the timing of when the LTV_{35} threshold was exceeded or the converted intensities exceeded 35 kt, the consensus intensity forecasts of the time that 35 kt is exceeded were evaluated and results were somewhat mixed. False alarms were improved upon somewhat (20 out of 33), but the misses showed only a slight improvement (two out of 20). It is evident that the uncertainty in the initial intensity value is not the only limitation in forecasting the time an intensity of 35 kt will be exceeded. The model forecasts of the relative vorticity and warm core also contribute to these timing forecast errors.

6. Adjusted Versus Unadjusted Forecast Improvement Points

A summary is given in Table 15 of the relative improvement due to the initial condition adjustments for the 850 – 500 mb relative vorticity and 700 – 500 mb warm core forecasts by the three global models and the consensus model of the time when the LTV_{35} or the intensity > 35 kt is first exceeded. In this somewhat arbitrary point scheme, two points are awarded for False Alarms or Misses that become Hits, but a penalty of two points is given if an unadjusted forecast Hit becomes a Miss when the initial condition adjustment is made. One-half point is awarded for a False Alarm that becomes a Miss since this at least provides the forecaster some information that the formation is yet to occur. Finally, one point is awarded (penalized) for an unadjusted Hit that remains as a Hit and provides a Better (Worse) forecast of the time of formation by at least 6 h.

First, comparing the improvements in the 850 – 500 mb relative vorticity among the three global models (Table 15), the NOGAPS was improved the smallest number of times from the initial adjustment to the CARQ intensity because so many of the adjusted forecasts were then Misses. This tendency indicates again that the NOGAPS relative vorticity forecasts have a systematic low bias. The improvements in the GFS and UKMO models are similar (Table 15) with the GFS adjustments providing slightly more improvements with the False Alarms, and slightly less improvement in the unadjusted

Misses category, than the UKMO adjustments provide. Both the GFS and UKMO model adjustments result in a number of penalty points for unadjusted Hits then becoming a Worse by at least 6 h or becoming a Miss.

Table 15. Improvement points for the adjusted versus the unadjusted 850 – 500 mb relative vorticity and 700 – 500 mb warm core forecasts by the three global models and the consensus (CON). The improvement points awarded for various categories of unadjusted forecasts are indicated in the left column, where False Alarms (FA), Hits (H), Misses (M), Better (B), and Worse (W). The sample sizes are 97, 95, 89, and 93 for the NOGAPS (N), GFS (G), UKMO (U), and consensus (CON) models, respectively.

Unadjusted To Adjusted	Award Points	Relative Vorticity			Warm Core			CON
		N	G	U	N	G	U	
FA → H	2.0	4	20	16	52	44	46	40
FA → M	0.5	9	5.5	5	7	6.5	9	6.5
H → B	1.0	3	7	4	5	13	8	5
H → W	-1.0	-2	-11	-7	-8	-17	-8	-5
H → M	-2.0	-8	-12	-10	-4	0	-4	-12
M → H	2.0	8	10	18	4	2	6	4
Total Points		14	19.5	26	56	48.5	57	38.5

The improvements in the 700 – 500 mb warm core (Table 15) due to the initial adjustment to the CARQ intensity are scored relatively higher than for the relative vorticity adjustments primarily because of the large numbers of False Alarms that become Hits, and the relatively small numbers of Misses. For these warm core adjustments to False Alarms, the NOGAPS model adjustments were scored slightly higher. The highest total score for initial adjustments was again for the UKMO model. The GFS model had a large number of unadjusted Hits that resulted in better timing of the time that the LTV warm core was exceeded, but had an even larger number of forecasts in which the timing was worse by at least 6 h.

The improvements in the consensus forecast due to the initial condition adjustments of both the relative vorticity and the warm core after conversion to equivalent intensities are intermediate between the separate adjustments (Table 15). This result is considered to be due to the degraded model performance for the relative vorticities. Consensus models are only better for combinations of skillful models. Thus, some improvement in the model-predicted tendency for relative vorticity is required.

B. RECOMMENDATIONS FOR FUTURE WORK

1. Sample Size Limitations

While 2005 was a record-setting year for tropical cyclone activity in the Atlantic, this study was still somewhat limited due to small sample sizes of tropical cyclone formation forecasts from each of the models. It is recommended that additional cases from the 2006 and 2007 seasons be examined to evaluate whether a more accurate representation of the LTVs of relative vorticity and warm core corresponding to CARQ intensities of 20 kt through 35 kt can be determined. Other combinations of variables calculated in the VORTRACK might be examined to see how well they might be related to tropical cyclone formation. For example, precipitation variables (both total and convective precipitation) that were available in NOGAPS and GFS models were not utilized by Cowan (2006) or in this study because they were absent for the UKMO model. These precipitation variables could be examined further since it seems reasonable that they would be related to how the tropical wave evolves as it intensifies into a tropical storm.

2. Application to Other Tropical Basins

It is also recommended that this post-processing methodology be tested for western North Pacific tropical cyclones to see what adjustments, if any, are needed for different tropical basins. Whereas the western North Pacific has more tropical cyclones and longer durations so that the sample size will be larger, the tropical cyclone formations are more related to the monsoon trough than to waves as in the Atlantic.

Perhaps this post-processing methodology is transportable in that it can be used in any basin, but it seems likely that the LTVs will be different for monsoon trough dominated regions.

3. Second Use of CARQ Intensities for Post-processing Model Forecasts

The first use of the CARQ intensities was to adjust the initial intensities of the tropical cyclones in the three numerical models to agree with the CARQ intensity via a post-processing technique. This initial condition adjustment of the 850 - 500 mb relative vorticities and the 700 - 500 mb warm core values led to some improvement in the adjusted forecasts versus the unadjusted forecasts of the time at which the Lower Tercile Value (LTV) of the relative vorticity or warm core corresponding to 35 kt would be exceeded. Converting the unadjusted and adjusted intensity forecasts and constructing consensus intensity from the three models and two variables again led to some improvement of the adjusted over the unadjusted consensus of when 35 kt would be exceeded.

A common conclusion from both the LTV_{35} exceedance test and the consensus intensity exceedance test was that the uncertainty in the forecast problem was not due only to the initial condition uncertainty. Rather, the model predictions of the relative vorticity tendency and the warm core tendency were also significant contributors to forecasts of when the LTV_{35} threshold or the 35 kt intensity would be exceeded. For example, the NOGAPS and UKMO relative vorticity tendencies had an obvious negative bias, and the NOGAPS warm core tendencies had a marked temporal oscillation that apparently included a diurnal effect. Temporal oscillations then led to excessive intensity oscillations when the conversions from predicted warm core values to intensity estimates were unbounded, and this degraded the consensus intensity forecasts. Evidence of these erroneous relative vorticity and warm core tendencies will be provided to the respective model developers in hopes that model modifications can be made in the future to reduce these biases.

A simple post-processing technique using the past trend in CARQ intensities is proposed as a first step in addressing the model forecast tendency contribution to the

intensity forecast uncertainty. This second use of the CARQ intensity trend follows a post-processing technique previously used by Elsberry and Frill (1980), Peak and Elsberry (1982), and by Ulses (1998) for initial motion vector errors in numerical weather prediction models. The basic assumption in this application is that the model intensity tendencies should be quasi-linear and should be related to persistence of past intensity trends, at least in the early life cycle of the tropical cyclone. That is, a backward extrapolation of the model intensity forecast tendencies (00 h-12 h, 00 h-24 h, 00 h-36 h; and perhaps 00 h-48 h) will be compared with the corresponding CARQ intensity values. The departures of the backward-extrapolated intensities and the CARQ intensities at -12 h, -24 h, -36 h, etc., then become one of the predictors in a statistical regression technique to adjust the model intensity forecasts at +12 h, +24 h, +36 h, etc. Depending on the weighting coefficient for the backward-extrapolated predictors, this post-processing of the model intensity forecasts may put a strong persistence of the intensity trend constraint on the model forecasts. In the track post-processing technique, the 00 h-12 h model forecast was strongly constrained to agree with persistence of past motion over 12 h, the 00 h-24 h model forecast was less constrained to agree with persistence, and the 00 h-36 h was minimally constrained to agree with persistence. A similar evolution of decreasing constraints is expected for this backward extrapolation of intensities to agree with the CARQ intensity trend.

Such a backward-extrapolation post-processing technique is expected to work well with the intensity forecasts converted from the 850 - 500 mb relative vorticities, because these relative vorticity errors seem to have a systematic low bias. However, the large volatility in the intensity values converted from the 700 - 500 mb warm core forecast values requires some conditioning to make these warm core forecasts tendencies self-consistent with the 850 - 500 mb relative vorticity tendencies. Part of this large volatility arises from the large intensity variations for only small changes in 700 - 500 mb warm core values in Figure 4. When this small slope is extrapolated to both larger and smaller intensities in applying the conversion of warm core forecast values to corresponding intensity forecast values, extreme variations in intensity occur for relatively small warm core changes. For the NOGAPS forecasts, strong convection that

leads to warming also leads to large convective momentum exchanges in the Emanuel convective parameterization technique, which reduces the tendency to spin-up the low-level vorticity and contributes to the low bias in intensity. If the convective momentum exchange coefficient is reduced, the NOGAPS has a tendency to spin-up false vortices. Since the specified convective momentum exchange coefficient is considered to be effective at stronger tropical cyclone stages, this undesirable tendency to under-estimate weaker stage cyclones is tolerated.

The first step is to put an upper bound on the intensity change values converted from the 700 - 500 mb warm core forecasts. That is, if the intensity change converted from the 850 - 500 mb relative vorticity is negative, then an upper bound of say 10 kt will be placed on the intensity increase converted from the 700 - 500 mb warm core forecast. This upper bound will damp the volatile intensity increases that resulted from the extrapolation of the lines in Figure 4 and its application to comparatively small warm core increases. Similarly the large negative intensity changes converted from the warm core decreases will be bounded to be no larger than -5 kt.

Perhaps a less empirical approach would be to attempt to estimate the slowly varying warm core trend by assuming that its large variability is primarily due to diurnal variability or to a false thermodynamical adjustment to the specified initial conditions. Thus, the six-hourly values of 700 - 500 mb warm core may be summed over 48 h (i.e., two diurnal cycles) and then the resulting sum would be linearly applied over the 48-h period. That is, one-quarter of the 48 h sum would be applied in the first 12 h and that warm core value would be converted with Figure 4 to a 12 h intensity value, and one half of the 48 h sum to the first 24 h, etc. Since these warm core changes would be slow, the upper and lower bounds would likely not be needed as the conversions to intensity would lie within the original (i.e., without extrapolation) intensity ranges of Figure 4. However, the use of average warm core trends may also result in too slow increases in intensity, and thus misses in formation forecasts. If so, some fraction of the warm core variability about the 48 h trend line might be retained (perhaps up to the intensity change bounds proposed above). The objective of the statistical regression post-processing technique is to adjust these warm core-derived intensity estimates to be self-consistent with the

relative vorticity-derived intensity estimates and also with the backward extrapolated predictors (and other predictors from the VORTRACK file) to obtain better intensity forecasts.

Whereas these proposed steps are empirical, they are an attempt to retain global numerical models as a tool in tropical cyclone intensity prediction for the early stages when other guidance is lacking or has deficiencies. These post-processing steps to address the contributions of model forecast uncertainty aspects to the intensity prediction problem in conjunction with addressing the initial intensity uncertainty is expected to improve the timing of when critical intensity thresholds associated with tropical cyclone formation will be exceeded.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A NOGAPS HISTOGRAMS

As indicated in Table 1, the CARQ intensity information was available for many hours prior to the initial time in the Best Track. These times with CARQ intensities were matched with the times at which the VORTRACK files had corresponding values of either 850 – 500 mb relative vorticity and 700 – 500 mb warm core values. The numbers of matches for the NOGAPS, GFS, and UKMO models were 131, 134, and 126 cases respectively.

Histograms for the relative vorticity and for the warm core corresponding to CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt were created. The numbers of cases are on the abscissa for each histogram. Ten bins of values of either 850 – 500 mb relative vorticity or 700 – 500 mb warm core are on the ordinate axis. Units for these variables are as follows: 850 – 500 mb relative vorticity is in units of 10^{-5} s^{-1} and 700 – 500 mb warm core is in degrees Celsius (°C). The histograms for the NOGAPS analyzed relative vorticities and warm cores are contained in this Appendix. Corresponding histograms for the GFS model and the UKMO model analyses are given in Appendices B and C respectively.

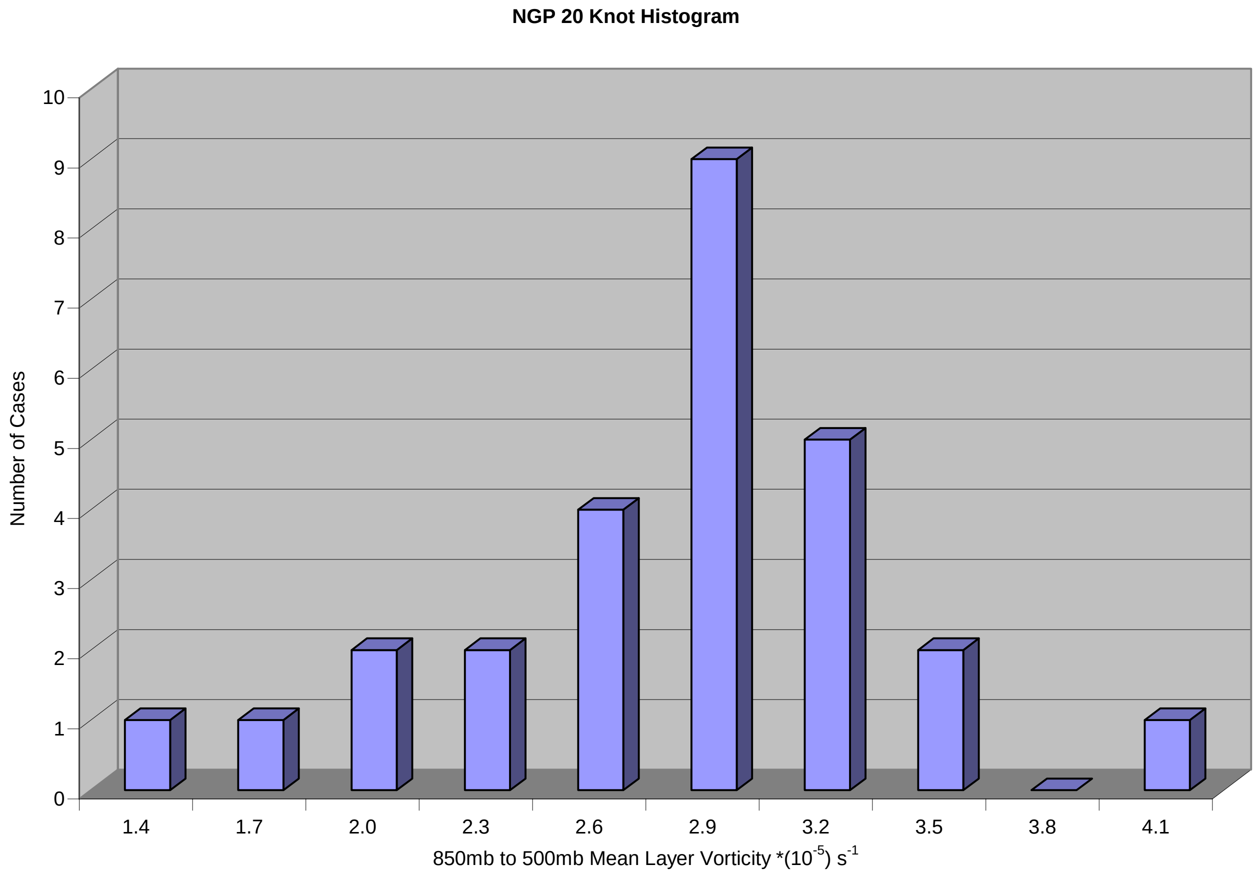


Figure A-1. NOGAPS 20 kt histogram for 850 - 500 mb relative vorticity

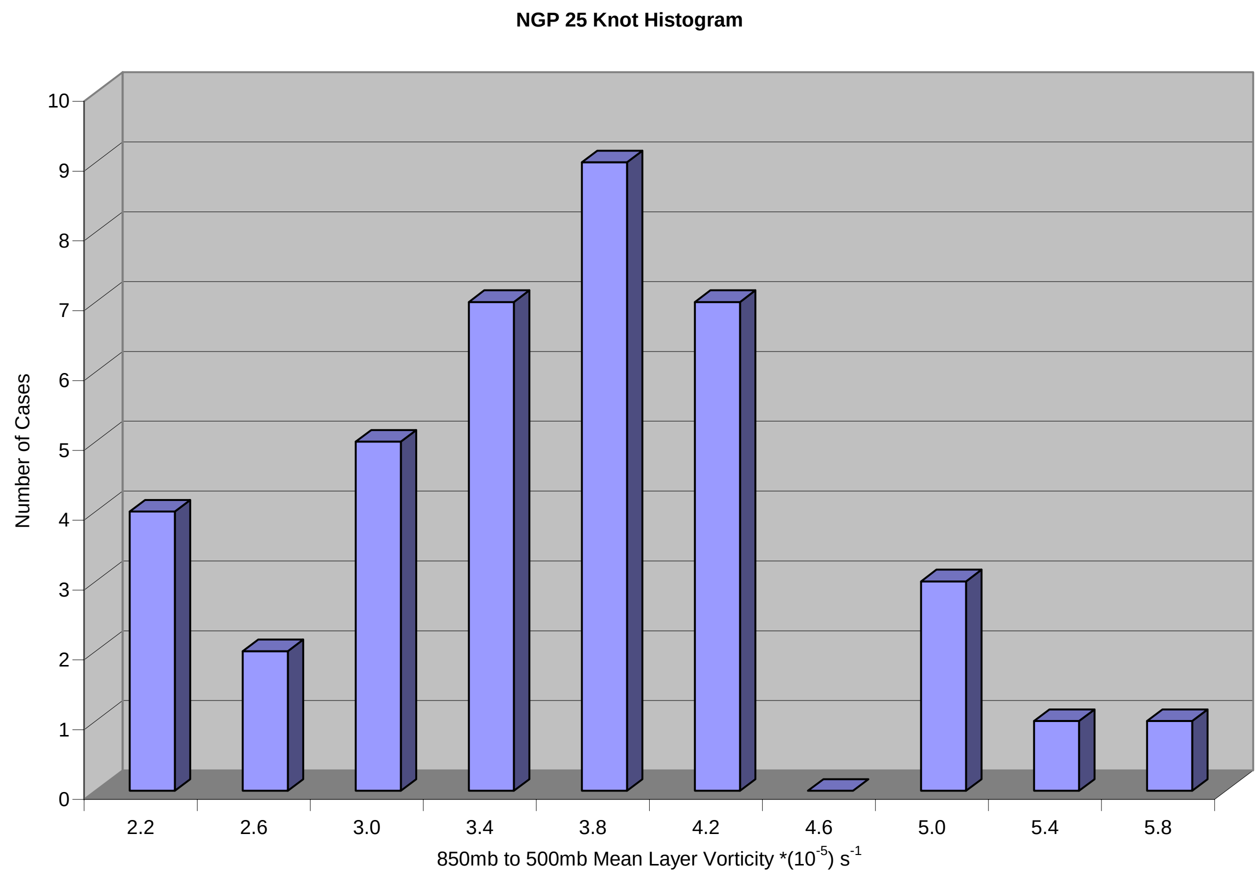


Figure A-2. NOGAPS 25 kt histogram for 850 - 500 mb relative vorticity

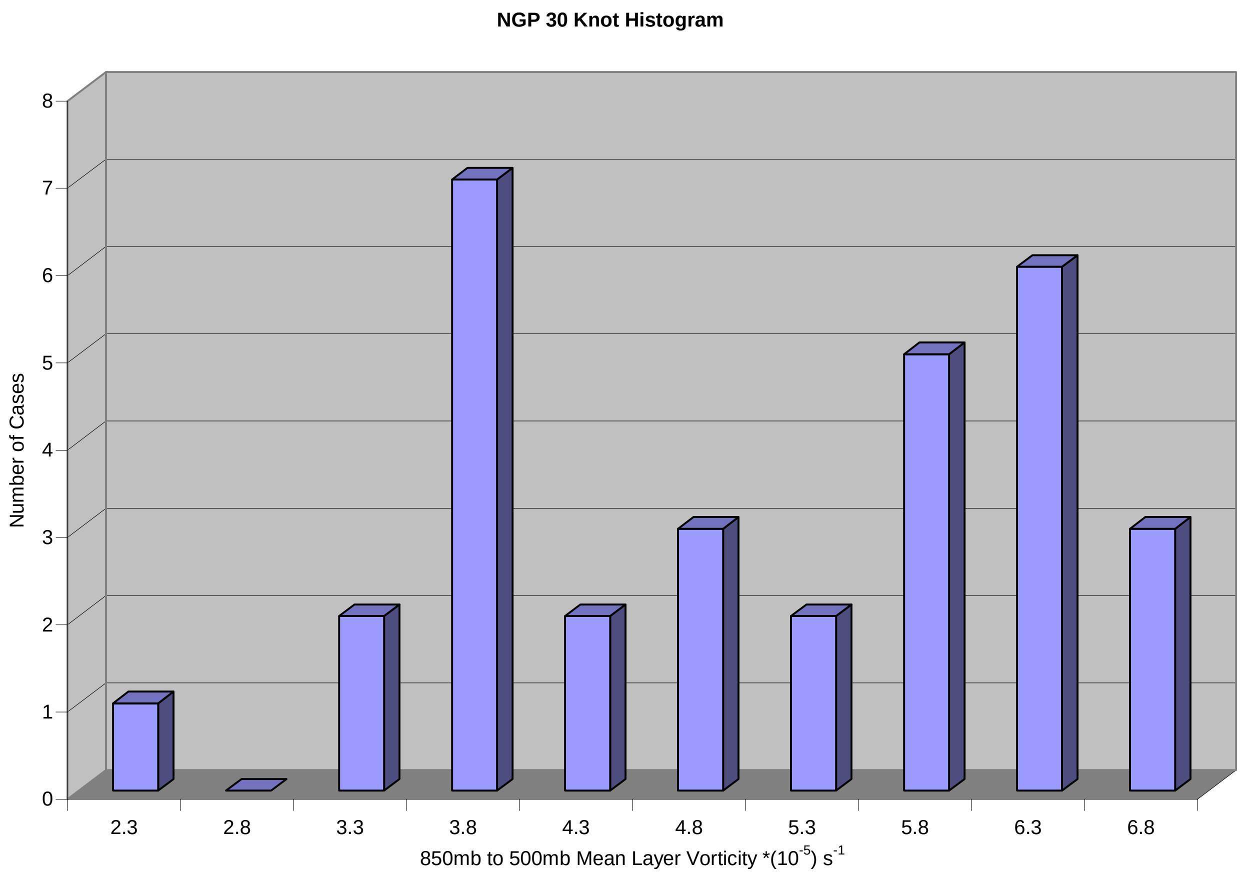


Figure A-3. NOGAPS 30 kt histogram for 850 - 500 mb relative vorticity

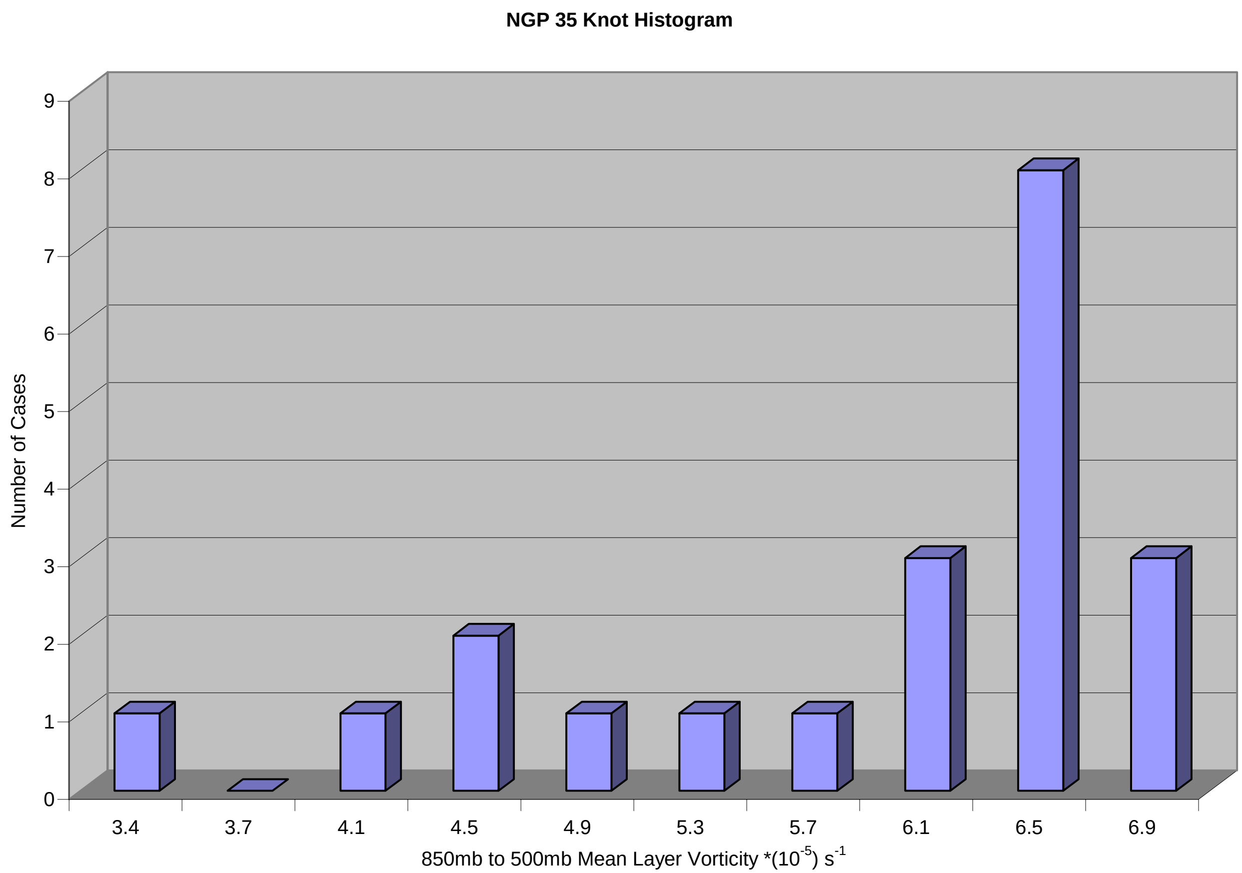


Figure A-4. NOGAPS 35 kt histogram for 850 - 500 mb relative vorticity

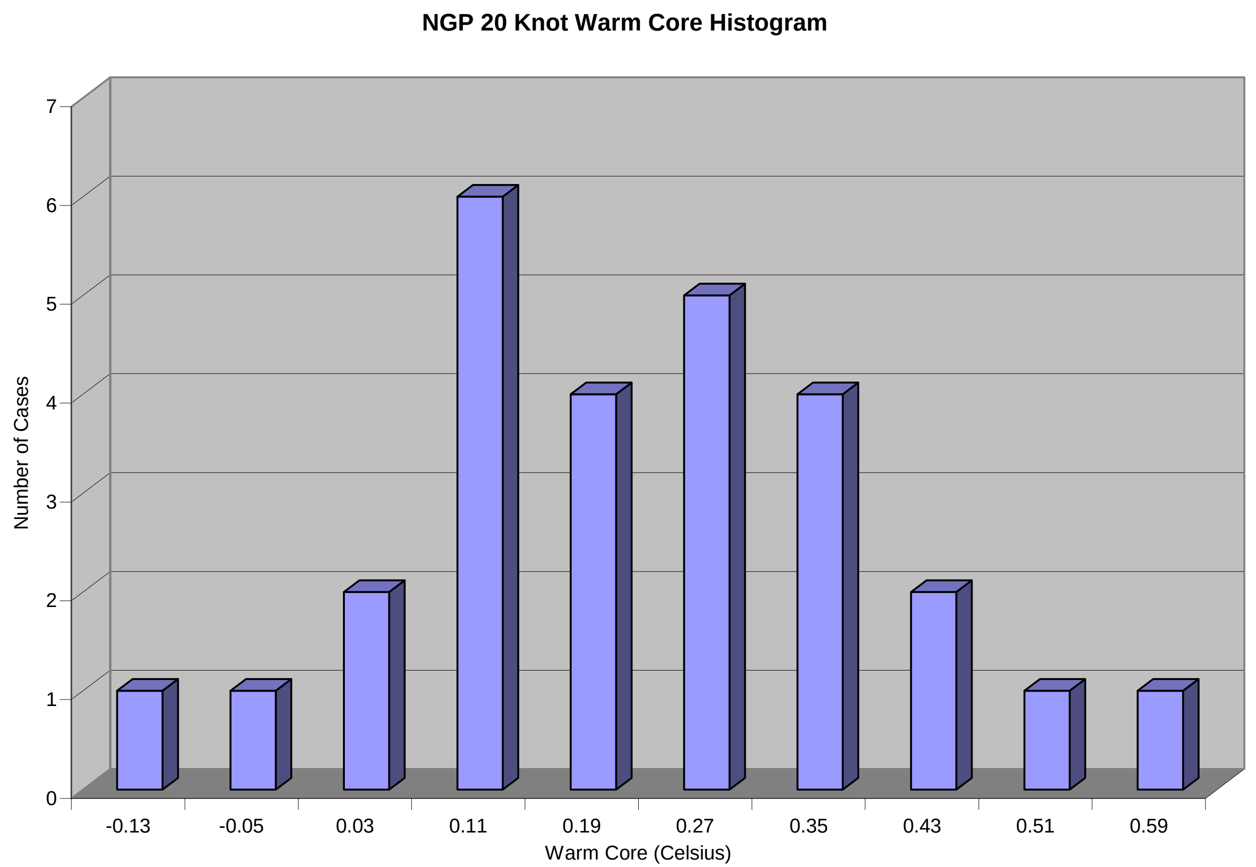


Figure A-5. NOGAPS 20 kt histogram for 700 - 500 mb warm core

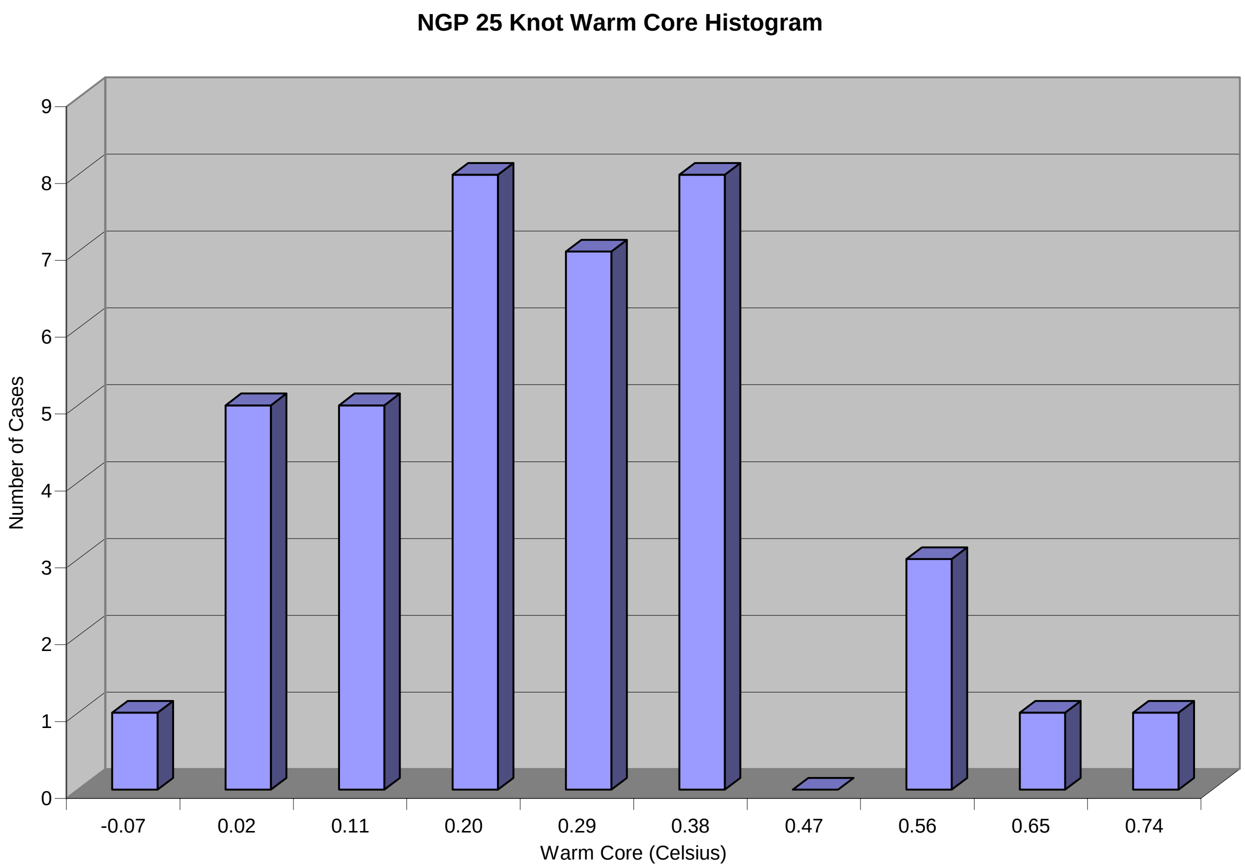


Figure A-6. NOGAPS 25 kt histogram for 700 - 500 mb warm core

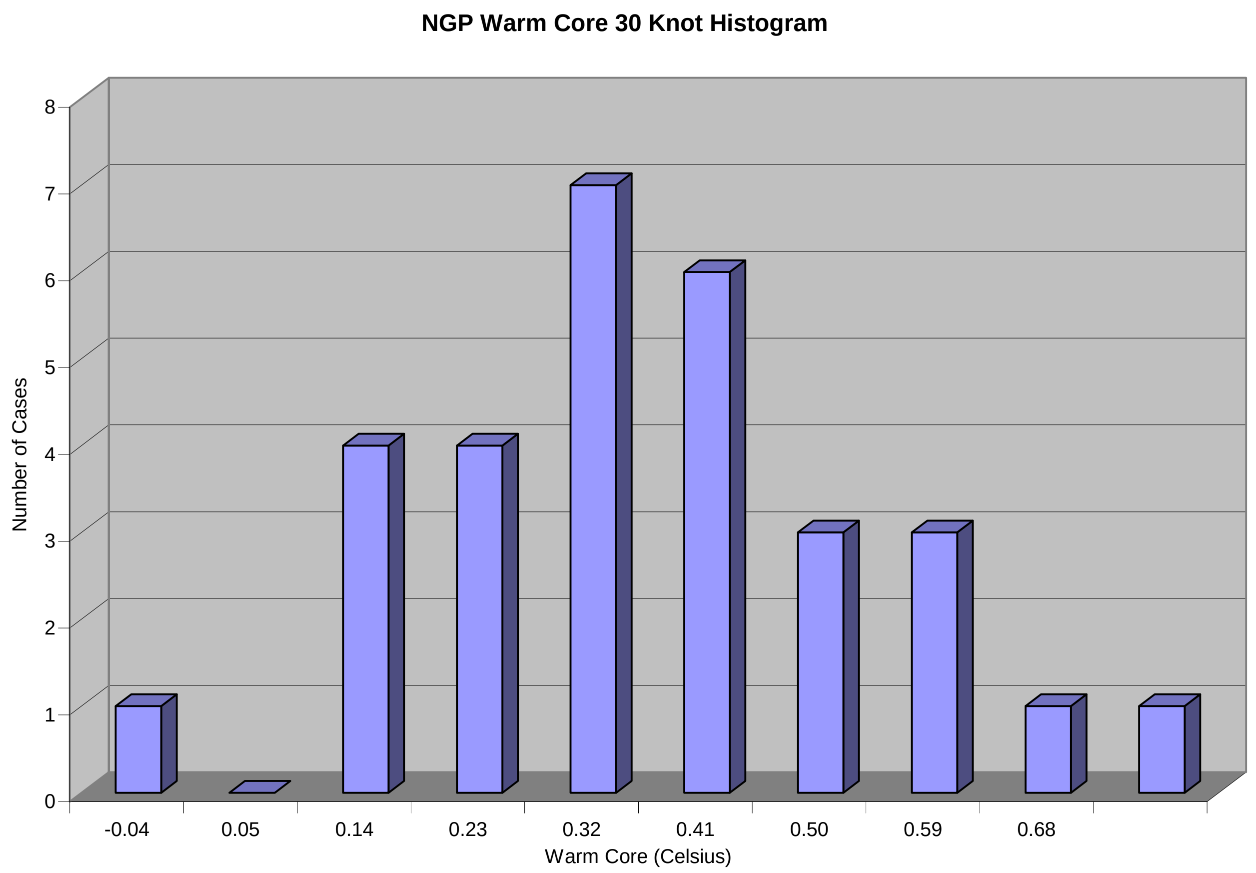


Figure A-7. NOGAPS 30 kt histogram for 700 - 500 mb warm core

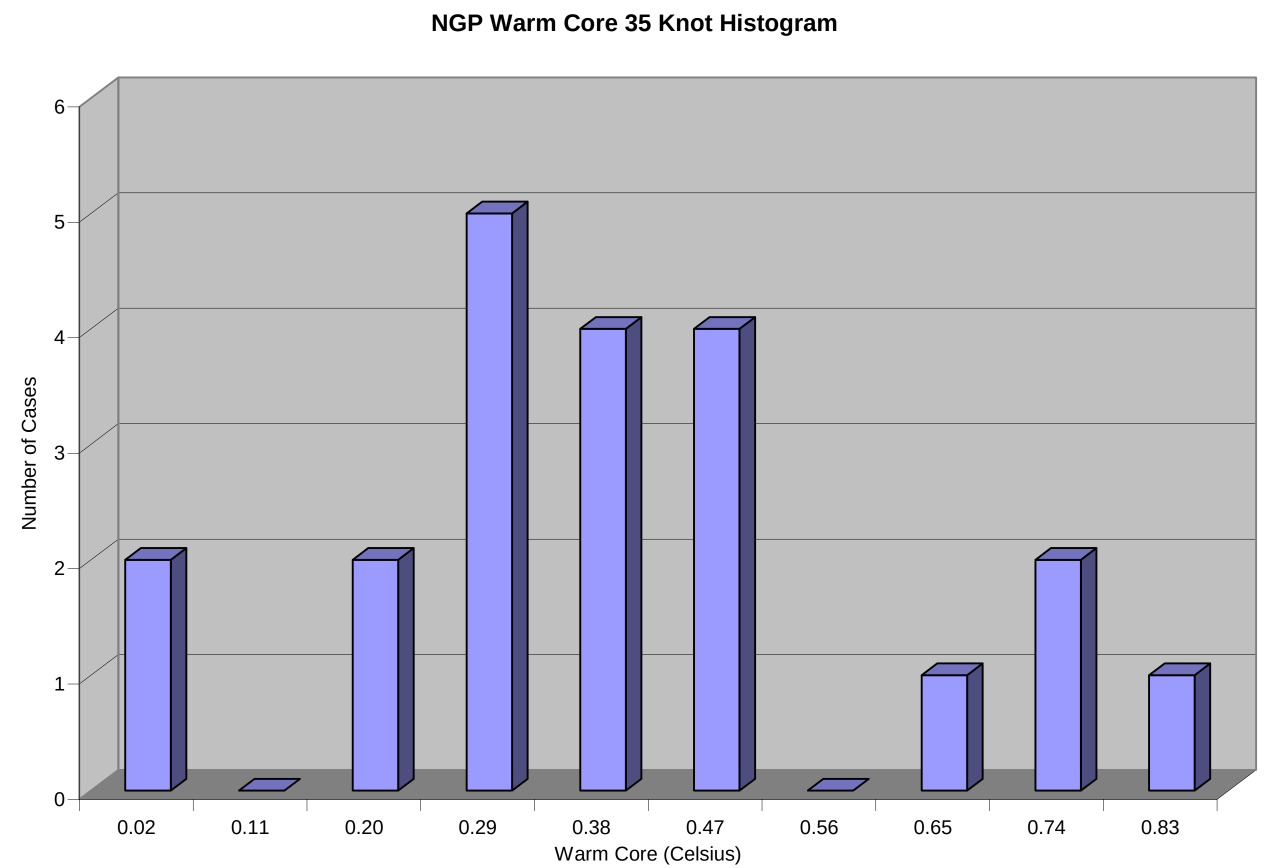


Figure A-8. NOGAPS 35 kt histogram for 700 - 500 mb warm core

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B GFS HISTOGRAMS

As in Appendix A, except for the GFS model analyses of 850 – 500 mb relative vorticity and 700 – 500 mb warm core values corresponding to CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt.

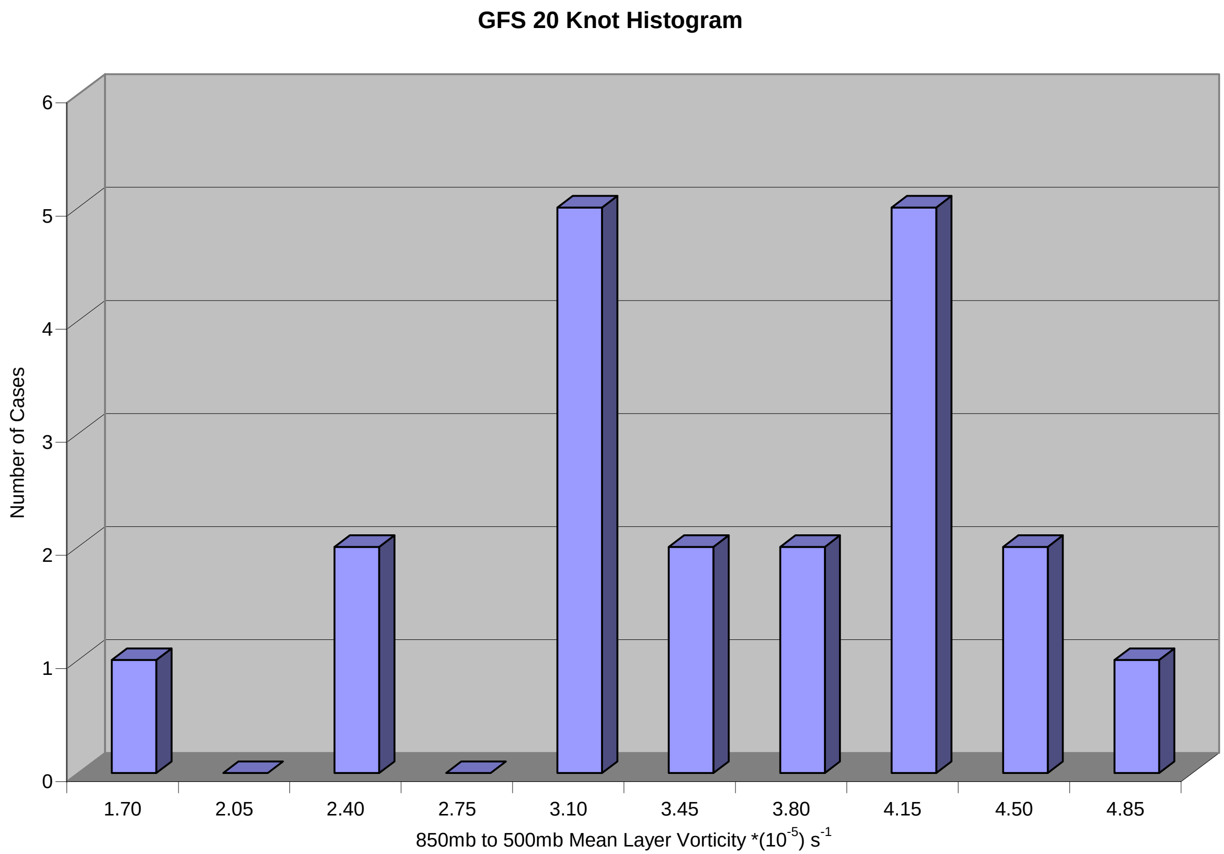


Figure B-1. GFS 20 kt histogram for 850 - 500 mb relative vorticity

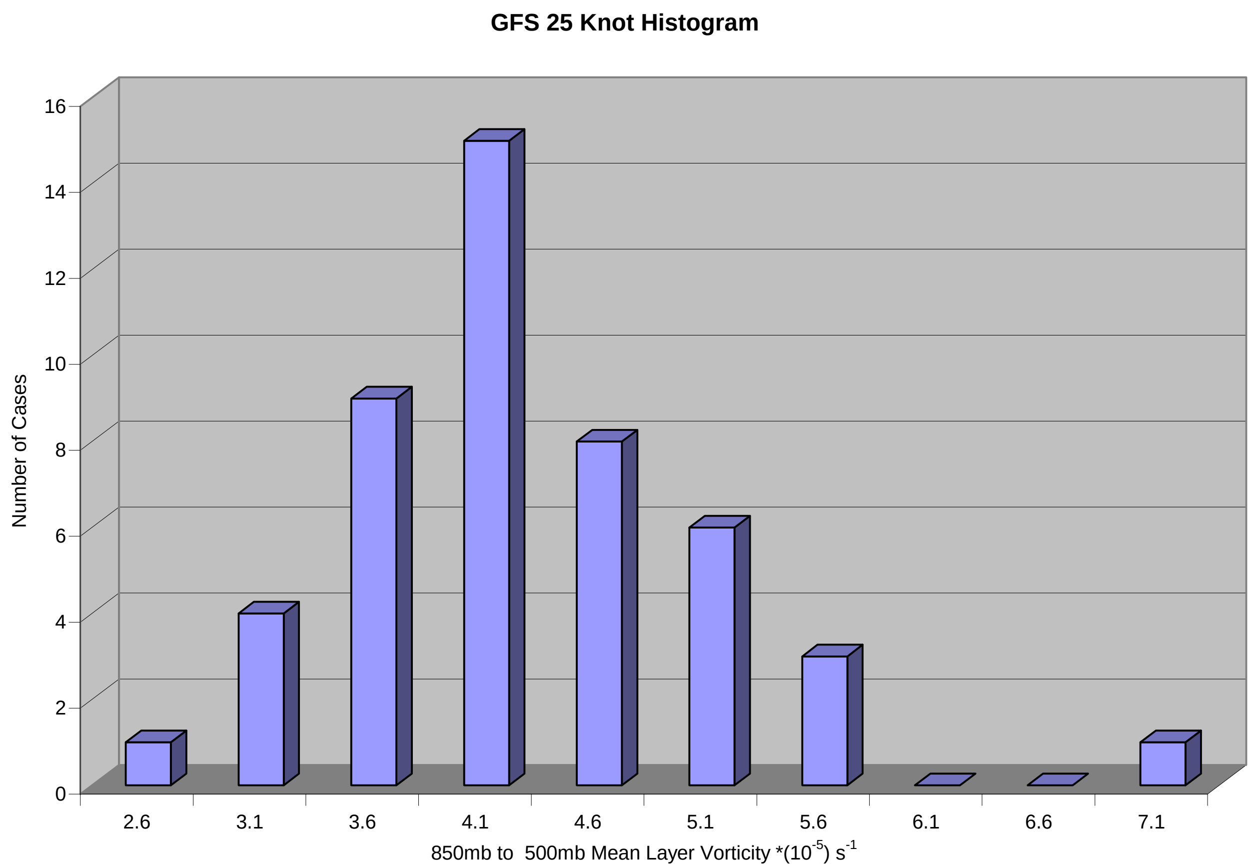


Figure B-2. GFS 25 kt histogram for 850 - 500 mb relative vorticity

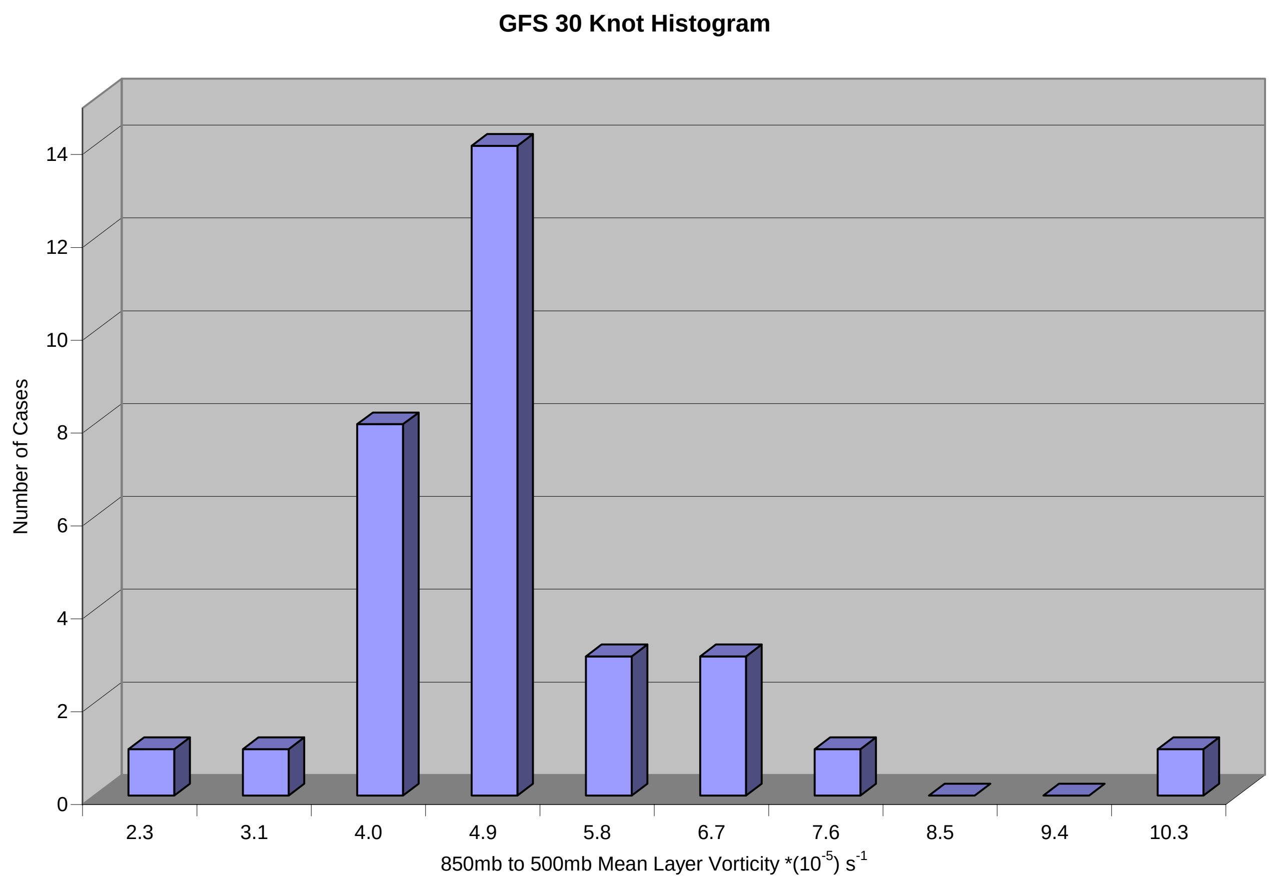


Figure B-3. GFS 30 kt histogram for 850 - 500 mb relative vorticity

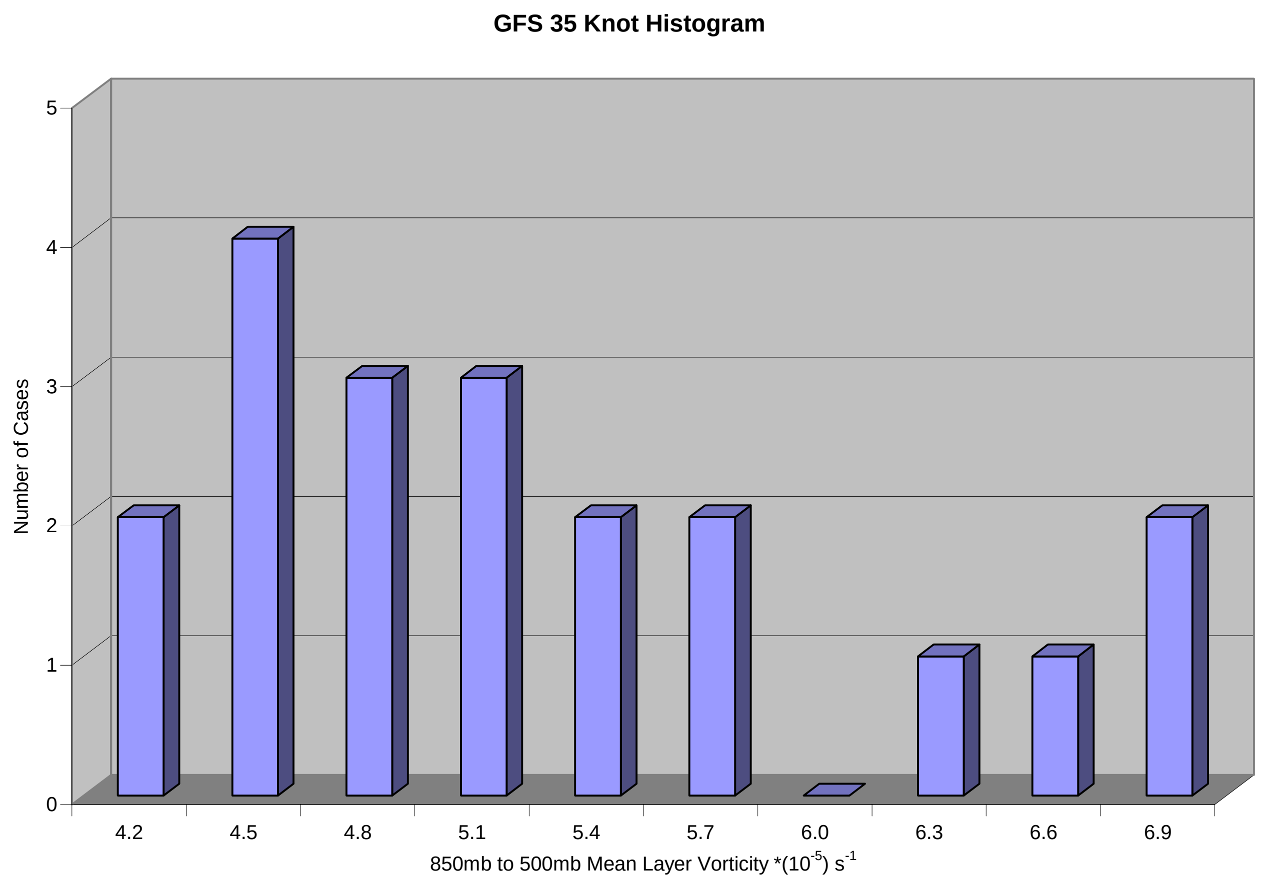


Figure B-4. GFS 35 kt histogram for 850 - 500 mb relative vorticity

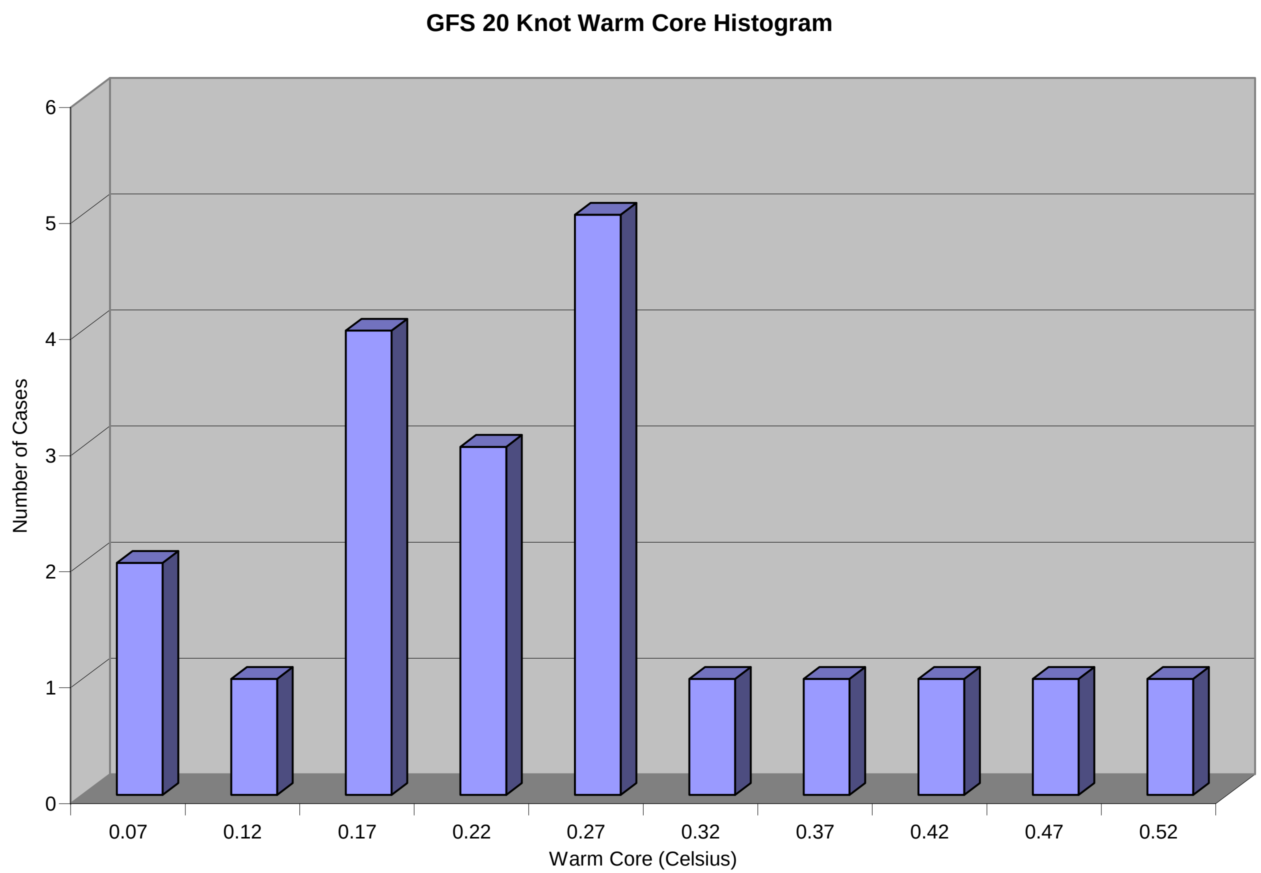


Figure B-5. GFS 20 kt histogram for 700 - 500 mb warm core

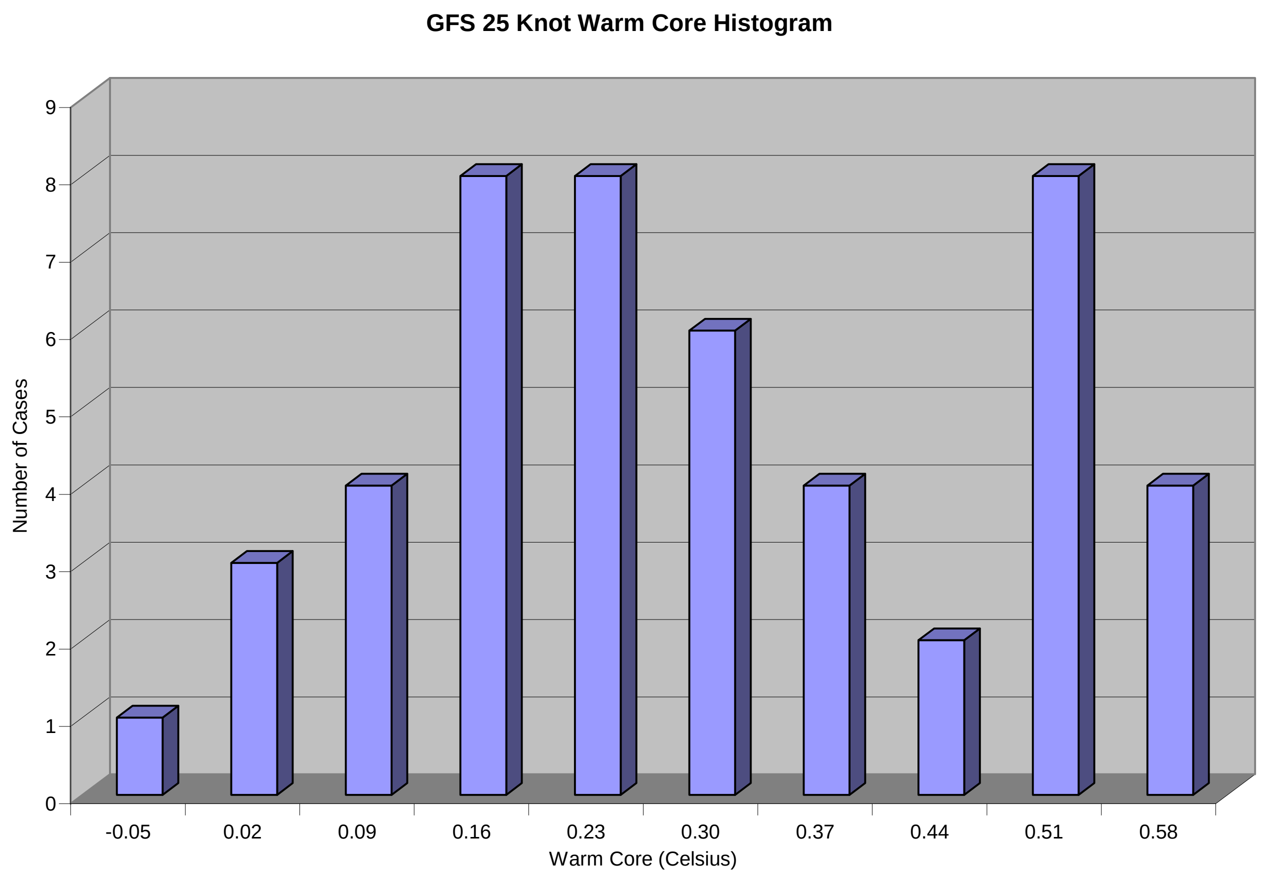


Figure B-6. GFS 25 kt histogram for 700 - 500 mb warm core

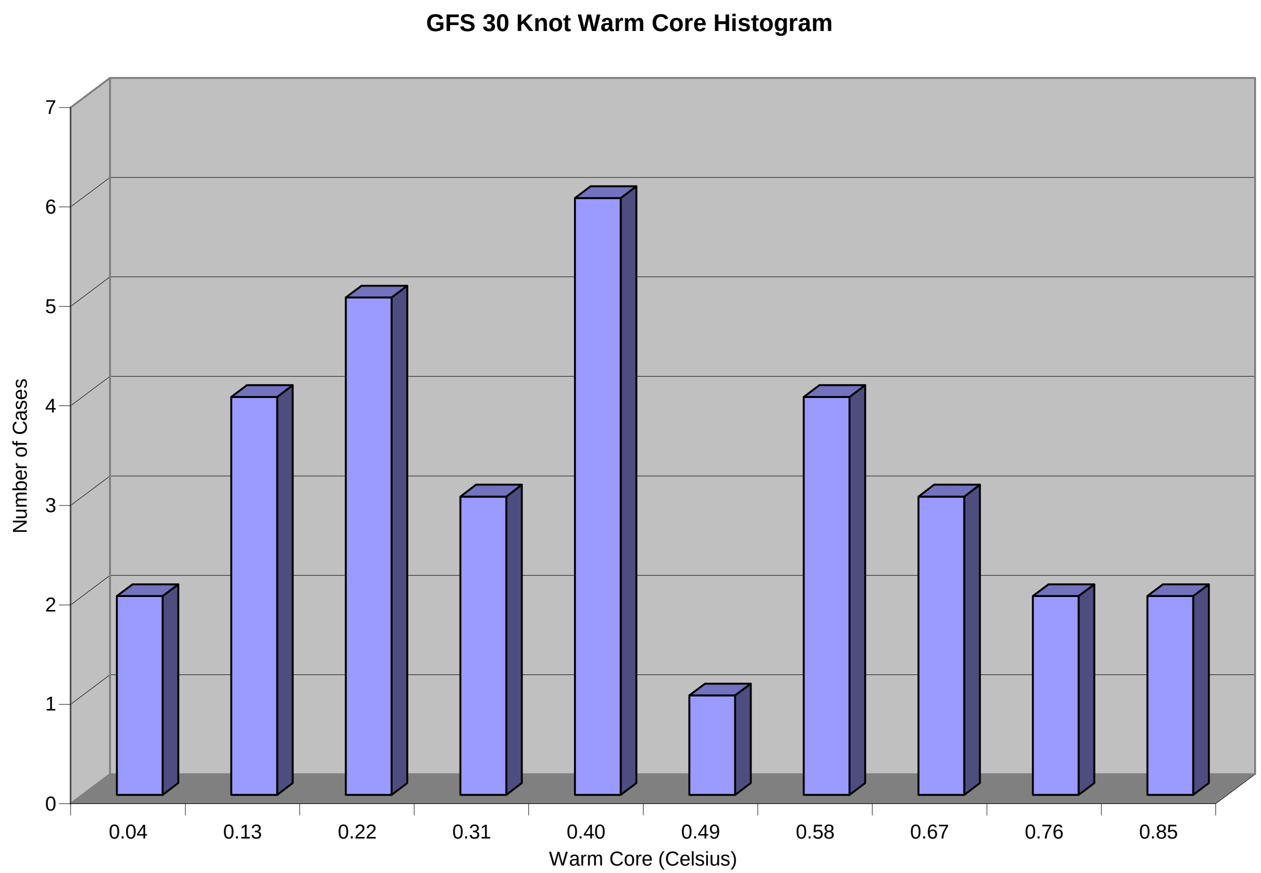


Figure B-7. GFS 30 kt histogram for 700 - 500 mb warm core

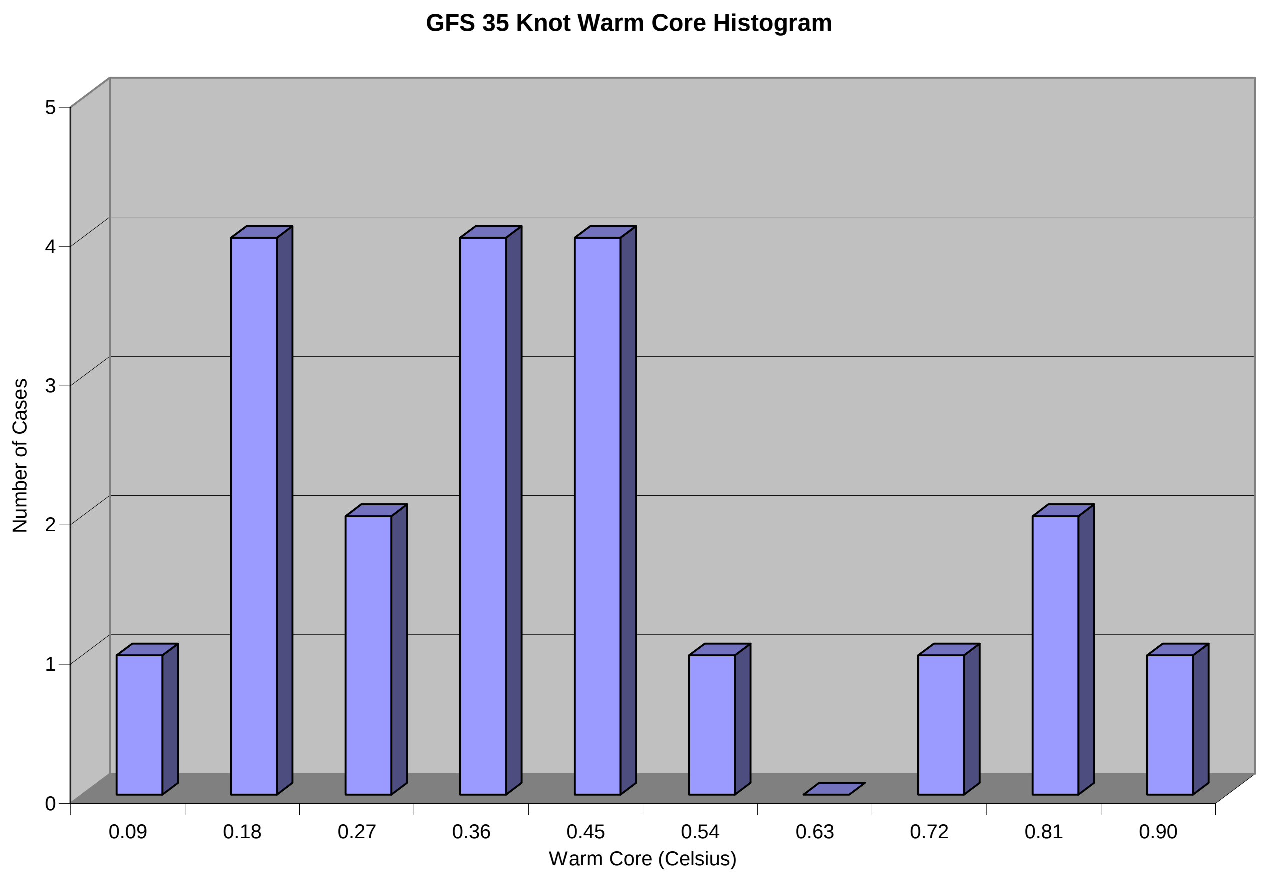


Figure B-8. GFS 35 kt histogram for 700 - 500 mb warm core

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C UKMO HISTOGRAMS

As in Appendix A, except for the UKMO model analyses of 850 – 500 mb relative vorticity and 700 – 500 mb warm core values corresponding to CARQ intensities of 20 kt, 25 kt, 30 kt, and 35 kt.

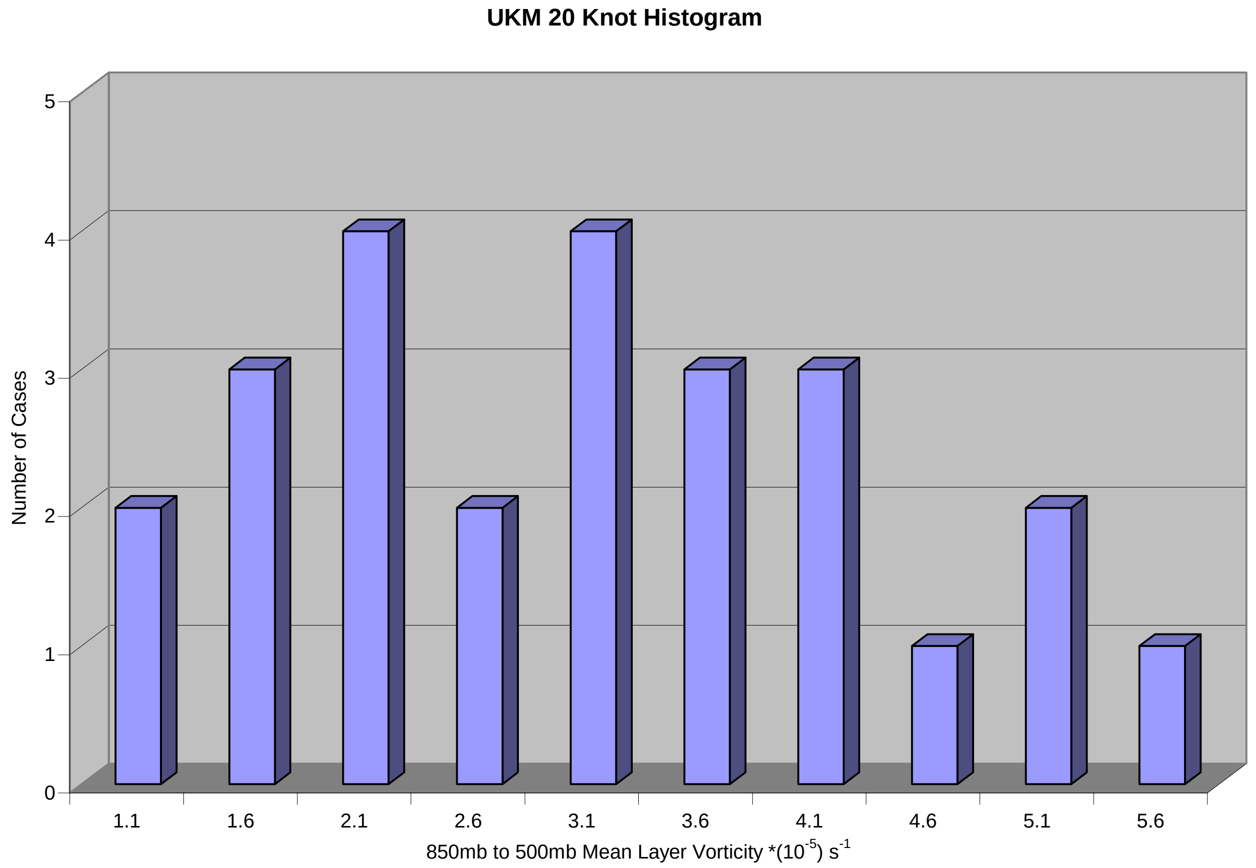


Figure C-1. UKMO 20 kt histogram for 850 - 500 mb relative vorticity

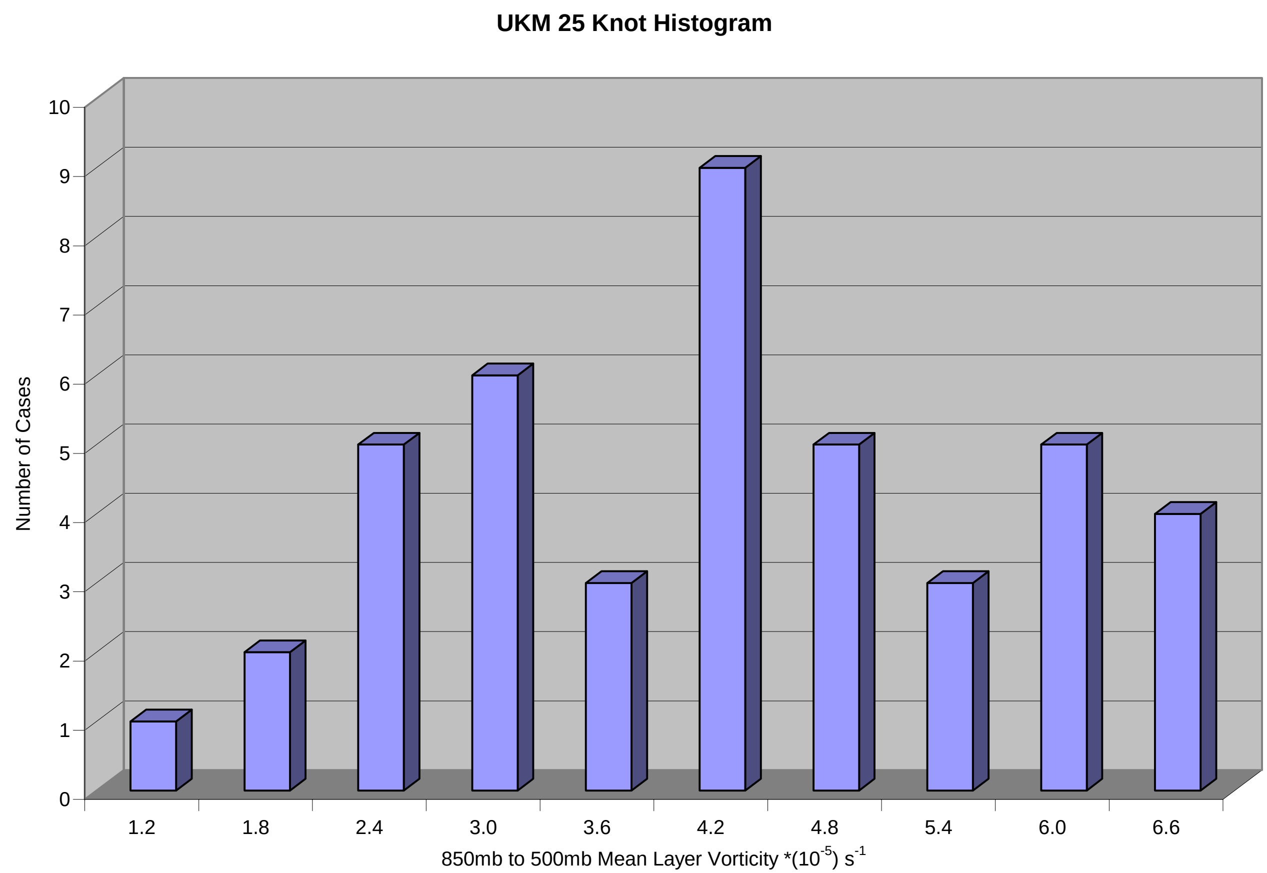


Figure C-2. UKMO25 kt histogram for 850 - 500 mb relative vorticity

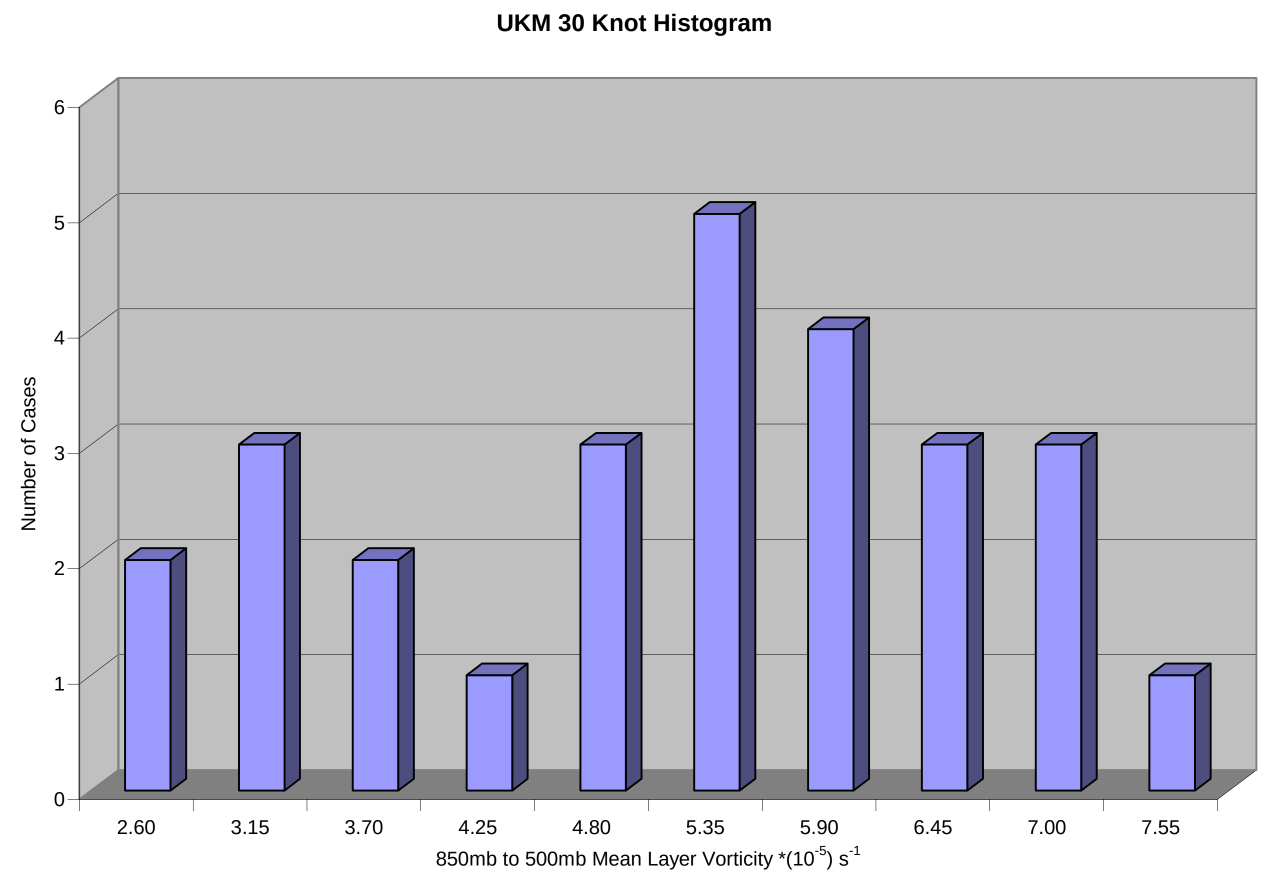


Figure C-3. UKMO 30 kt histogram for 850 - 500 mb relative vorticity

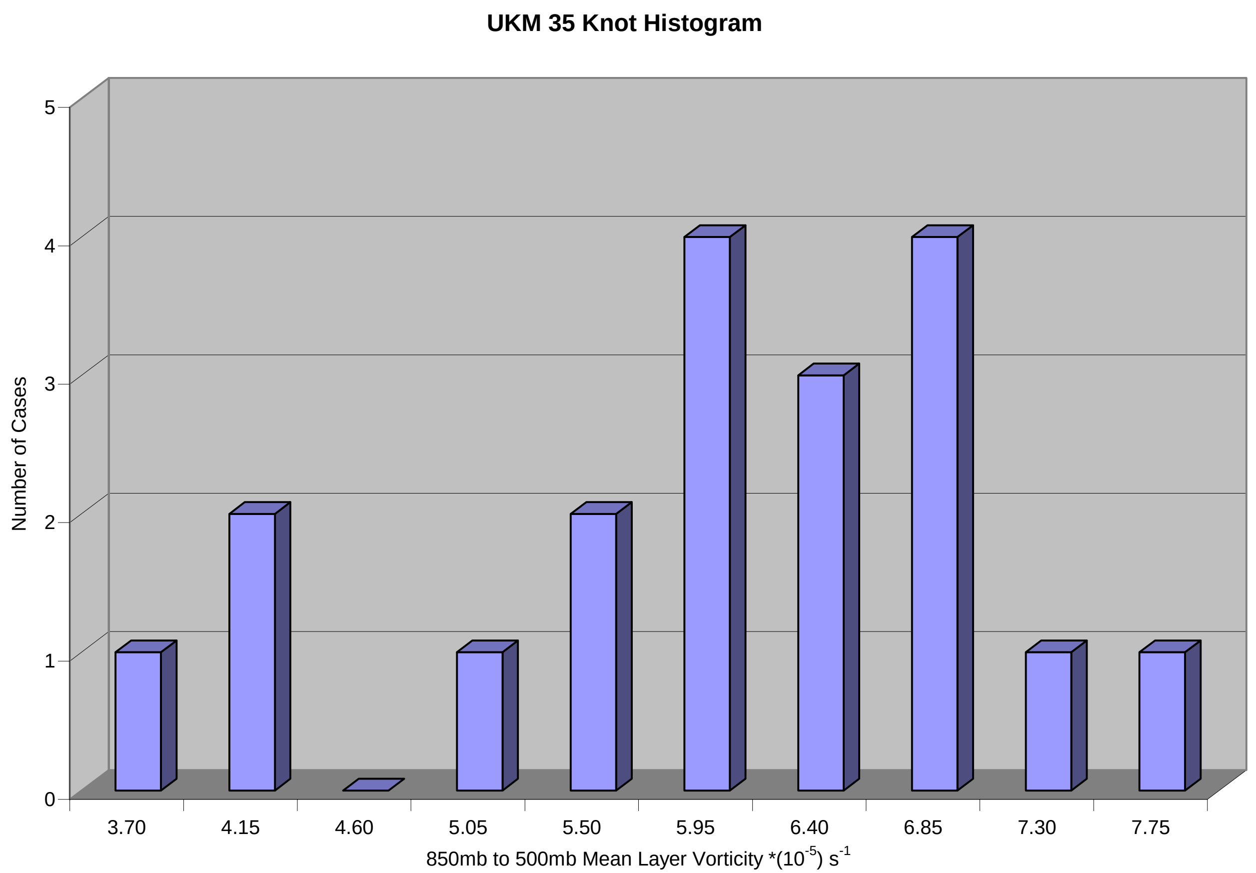


Figure C-4. UKMO 35 kt histogram for 850 - 500 mb relative vorticity

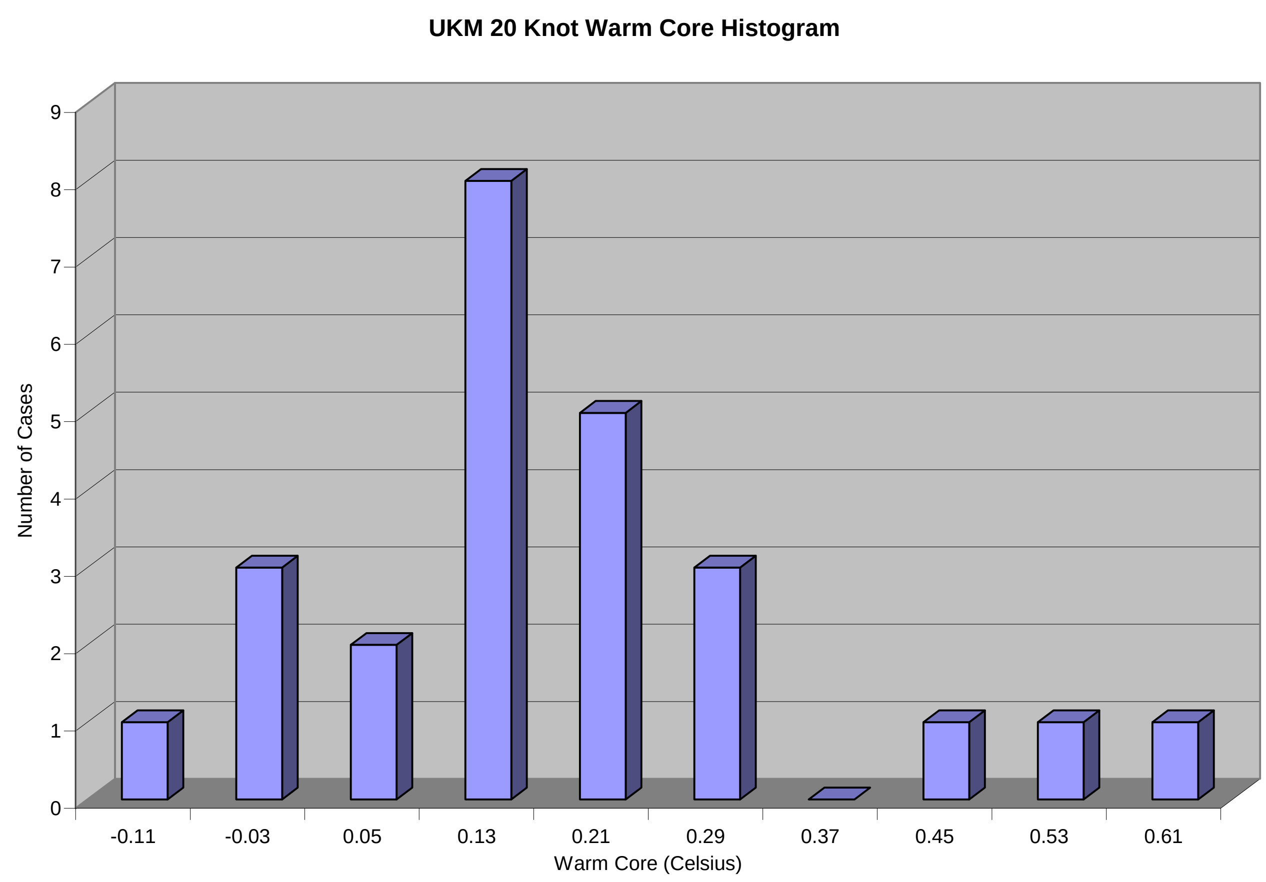


Figure C-5. UKMO 20 kt histogram for 700 - 500 mb warm core

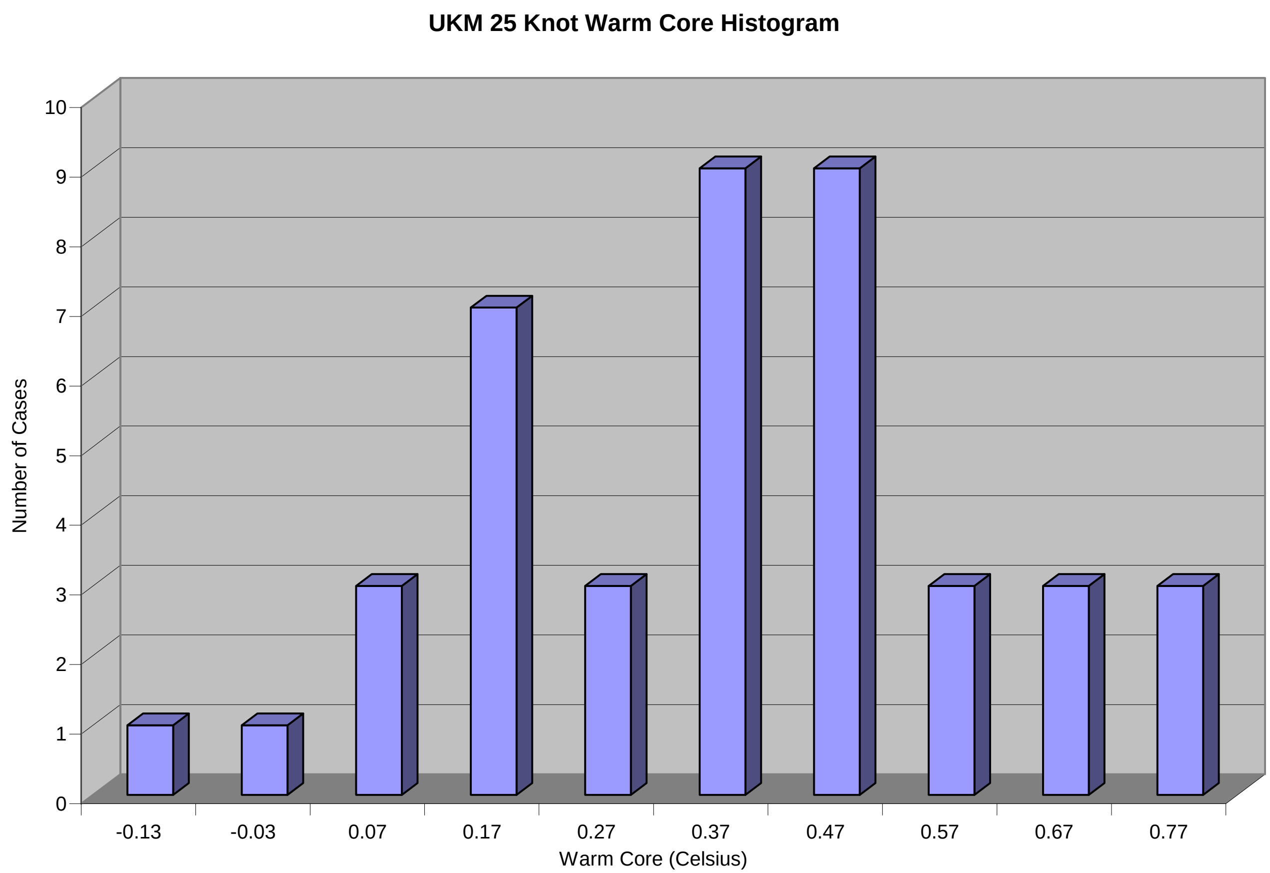


Figure C-6. UKMO 25 kt histogram for 700 - 500 mb warm core

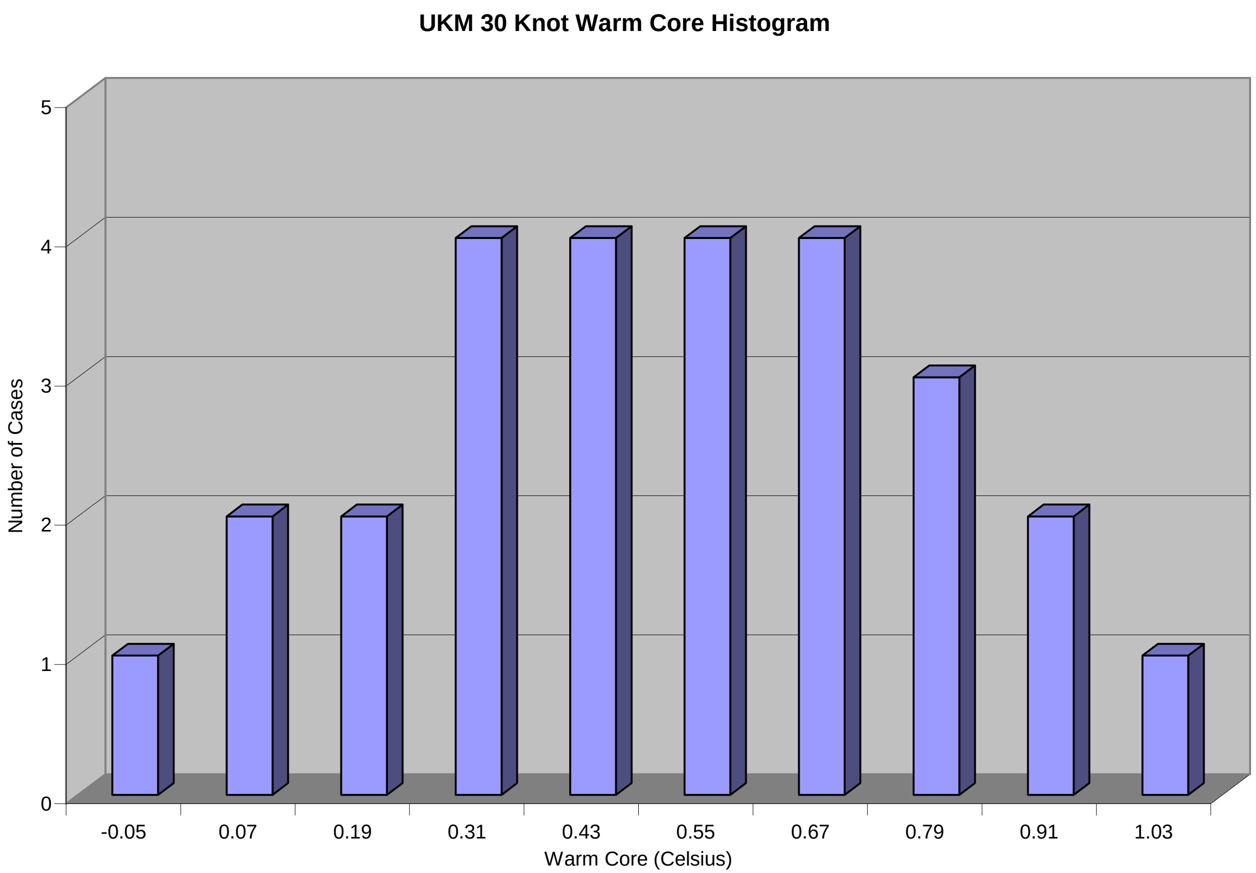


Figure C-7. UKMO 30 kt histogram for 700 - 500 mb warm core

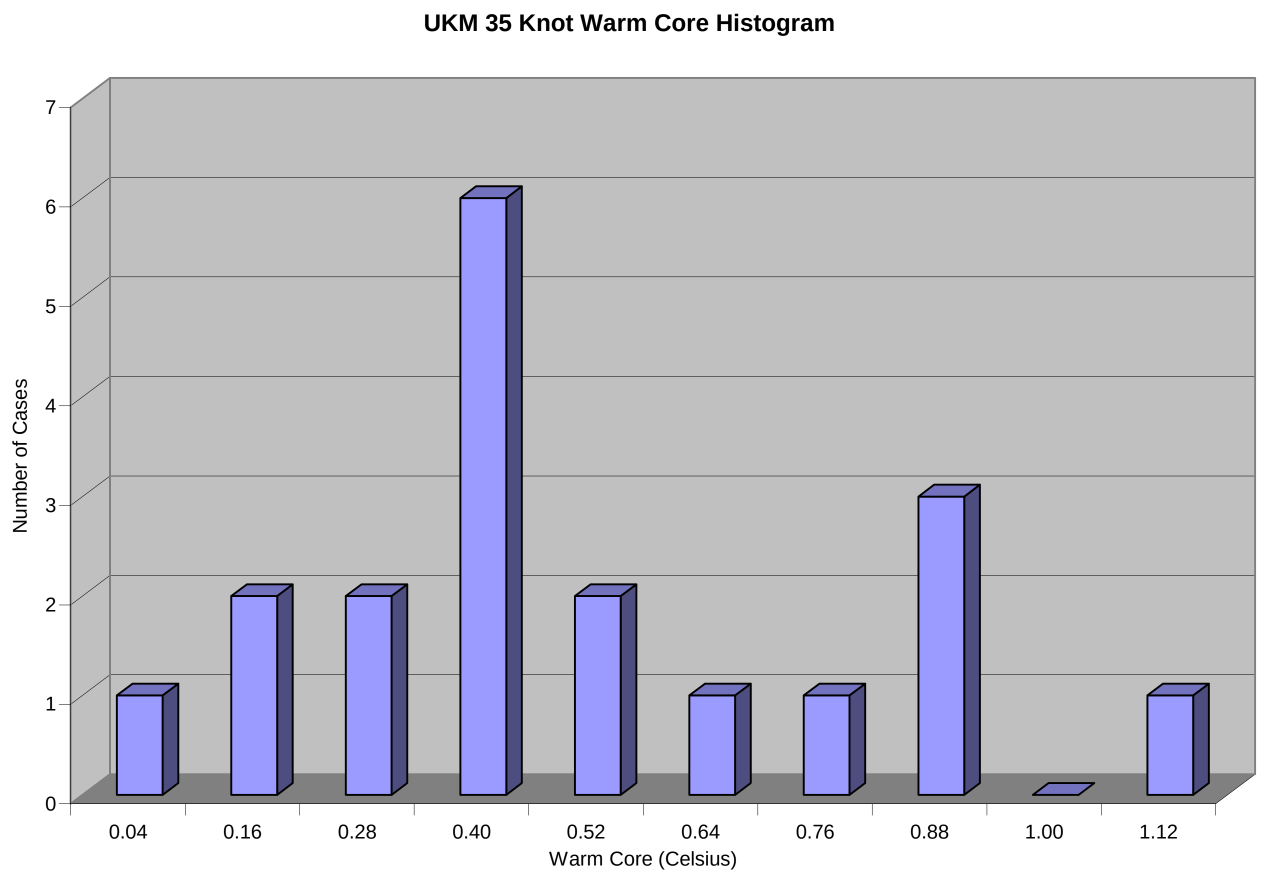


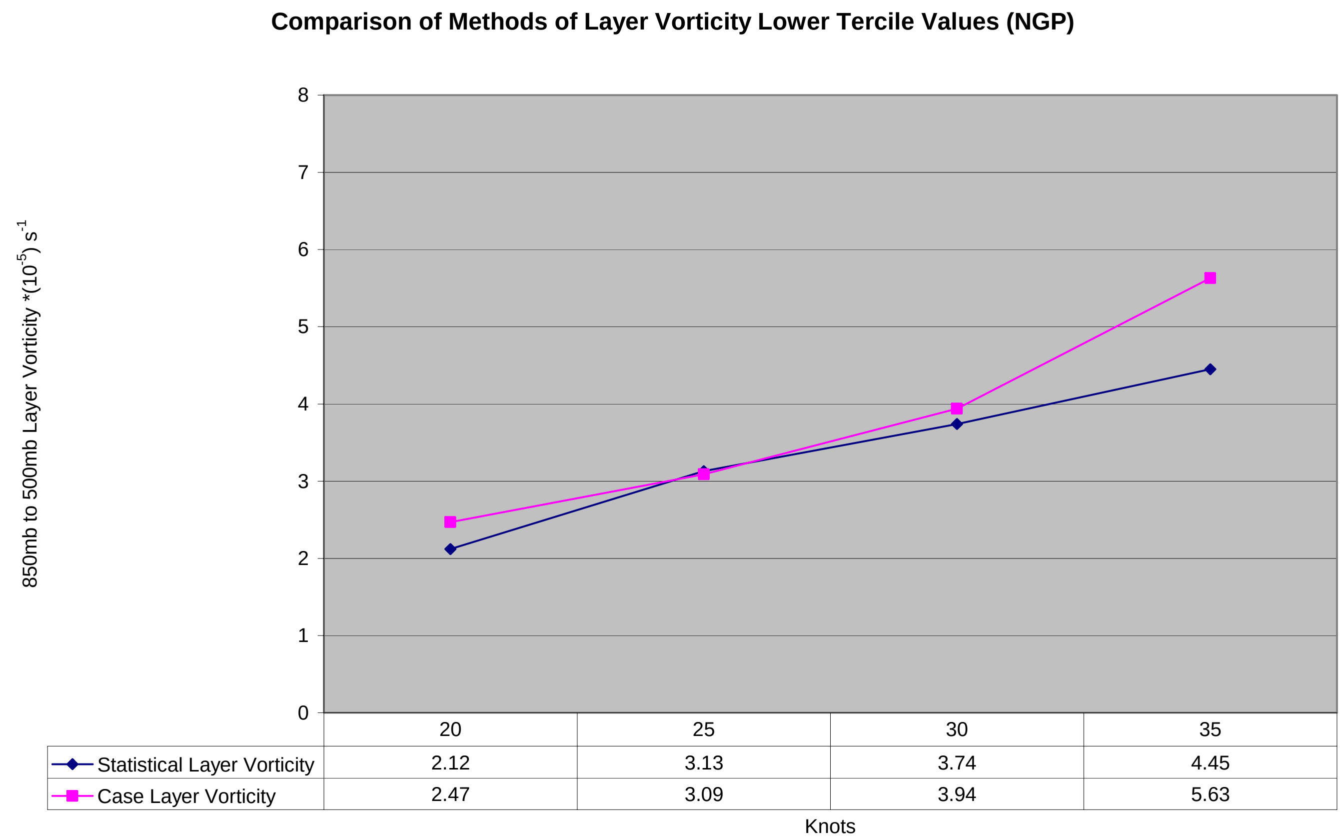
Figure C-8. UKMO 35 kt histogram for 700 - 500 mb warm core

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D LOWER TERCILE VALUES

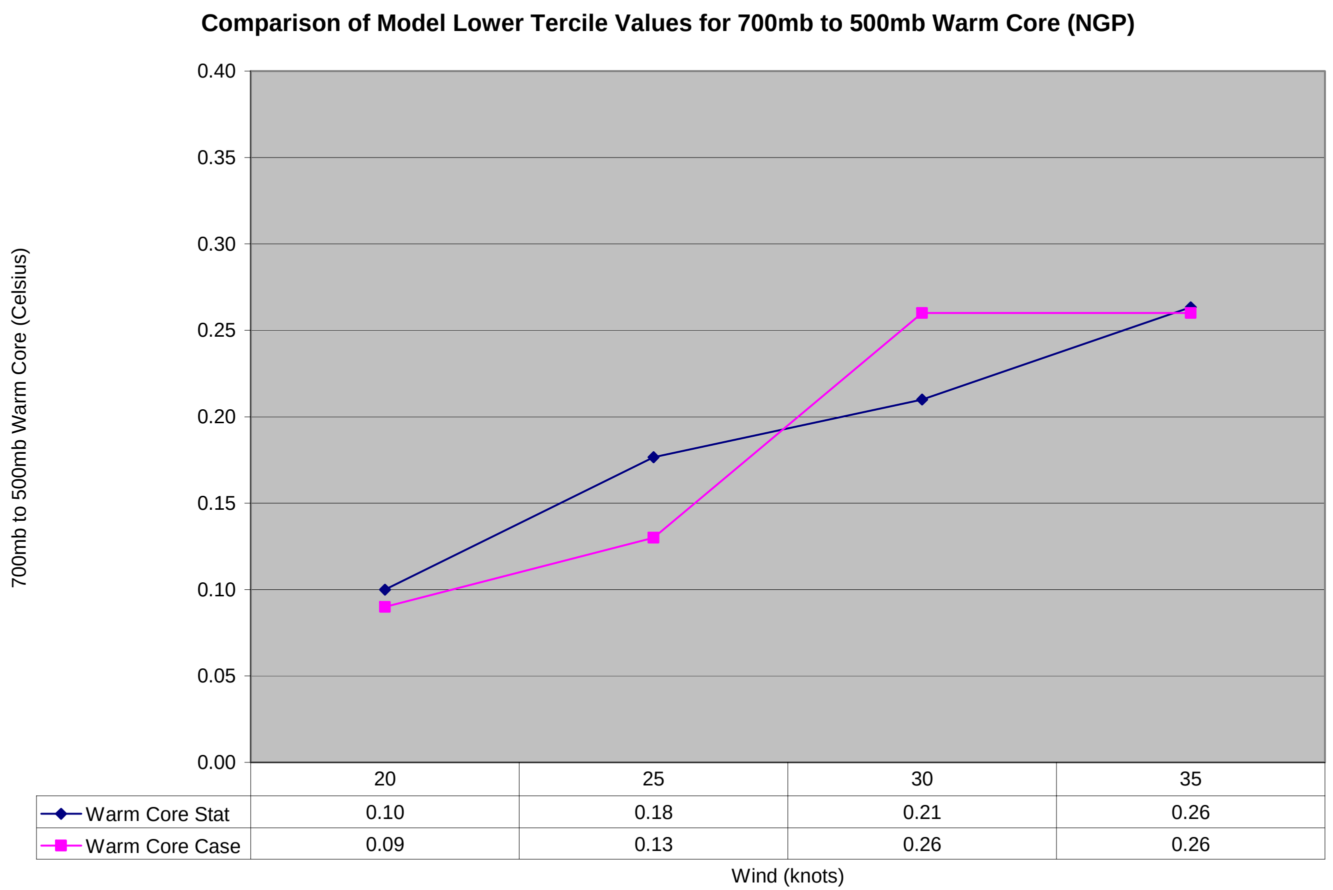
Lower Tercile Values (LTVs) of 850 - 500 mb relative vorticity and 700 – 500 mb warm core for all three models (NOGAPS, GFS, and UKMO) corresponding to CARQ intensities of 20 kt, 25 kt, 30kt, and 35 kt were derived for use as adjustments to the initial conditions in the post-processing technique (Figure 1) and for use as the 35 kt threshold as formation. Units for these variables are: 850 – 500 mb relative vorticity is in units of 10^{-5} s^{-1} and 700 – 500 mb warm core is in degrees Celsius (°C). As described in Chapter III. A.1., Lower Tercile Values (LTVs) for each of the histograms in Appendices A (NOGAPS), B (GFS), and C (UKMO) were derived by two methods: statistical fit technique and counting one third of the cases (labeled Case on the following diagrams). These distributions are plotted as line graphs with symbols as denoted in the inset in the lower right corner.

On the following pages, two figures comparing LTVs derived by the case and statistical methods for the NOGAPS 850 - 500 mb relative vorticity and 700 – 500 mb warm core will be presented, then two for GFS, and two for the UKMO model. In the final section of this appendix, separate figures comparing the LTVs for the three models for 850 - 500 mb relative vorticity at 20, 25, 30, and 35 kt derived using case method and then derived using the statistical methods will be presented. Finally, two figures comparing the LTVs of all three models LTV for 700 - 500 mb warm core at 20, 25, 30, and 35 kt using the case method first and then the statistical method will be presented.



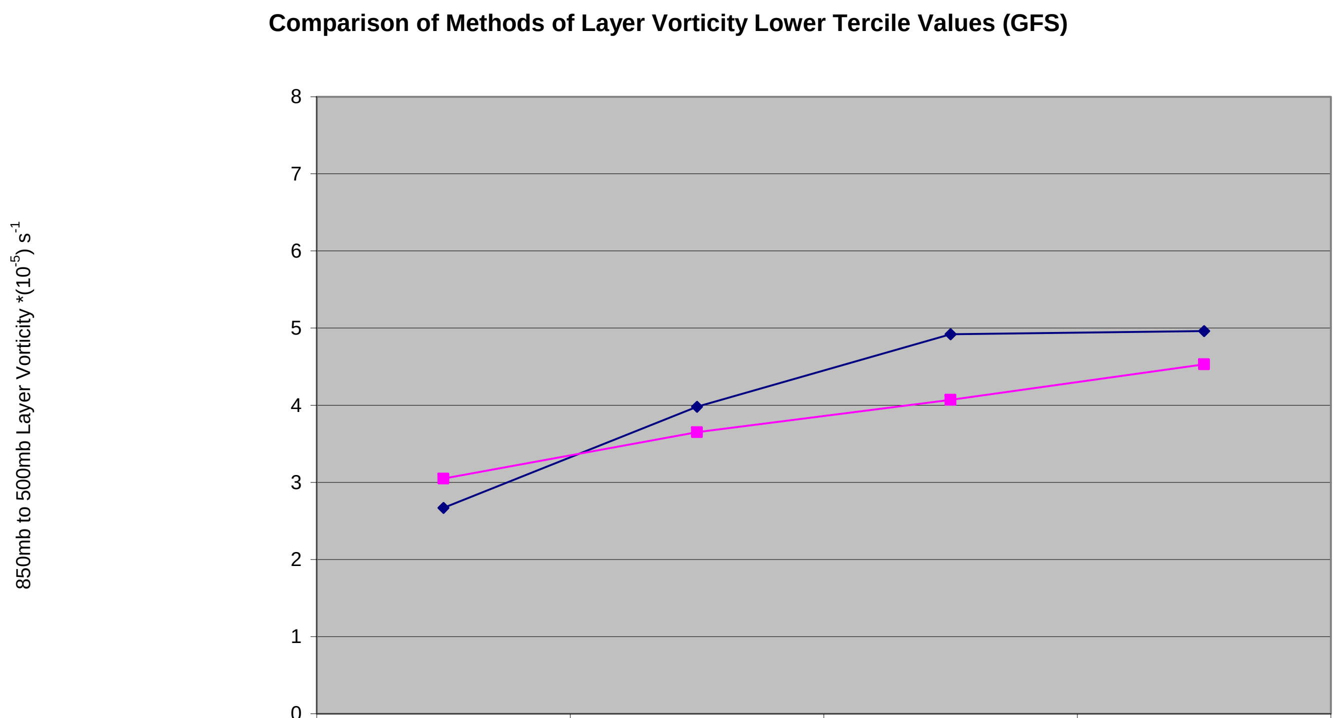
	20	25	30	35
Statistical Layer Vorticity	2.12	3.13	3.74	4.45
Case Layer Vorticity	2.47	3.09	3.94	5.63

Figure D-1. NOGAPS LTV for case and statistical methods of calculating 850 - 500 mb relative vorticity for 20, 25, 30, and 35 kt.



	20	25	30	35
Warm Core Stat	0.10	0.18	0.21	0.26
Warm Core Case	0.09	0.13	0.26	0.26

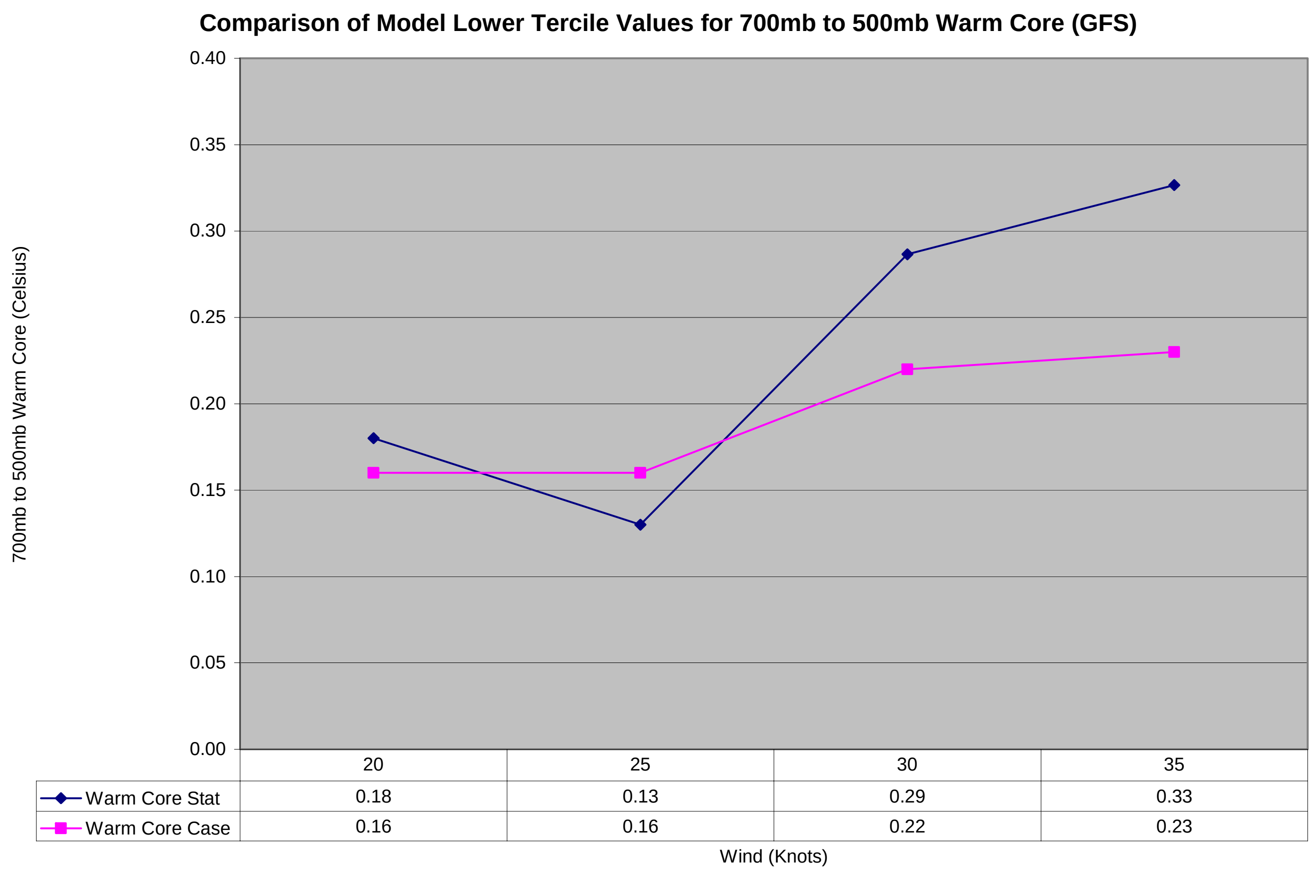
Figure D-2. NOGAPS LTV for case and statistical methods of calculating 700 - 500 mb warm core for 20, 25, 30, and 35 kt



	20	25	30	35
Statistical Layer Vorticity	2.67	3.98	4.92	4.96
Case Layer Vorticity	3.05	3.65	4.07	4.53

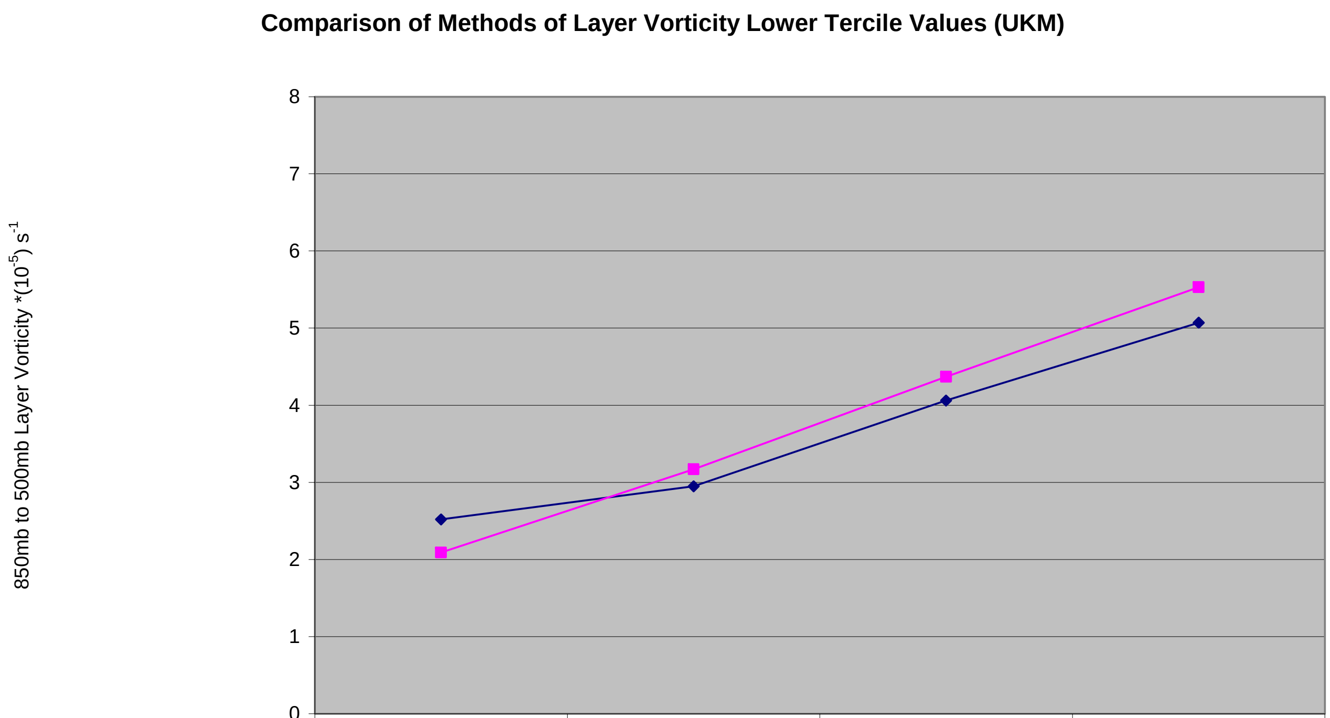
Knots

Figure D-3. GFS LTV for case and statistical methods of calculating 850 - 500 mb relative vorticity for 20, 25, 30, and 35 kt



	20	25	30	35
Warm Core Stat	0.18	0.13	0.29	0.33
Warm Core Case	0.16	0.16	0.22	0.23

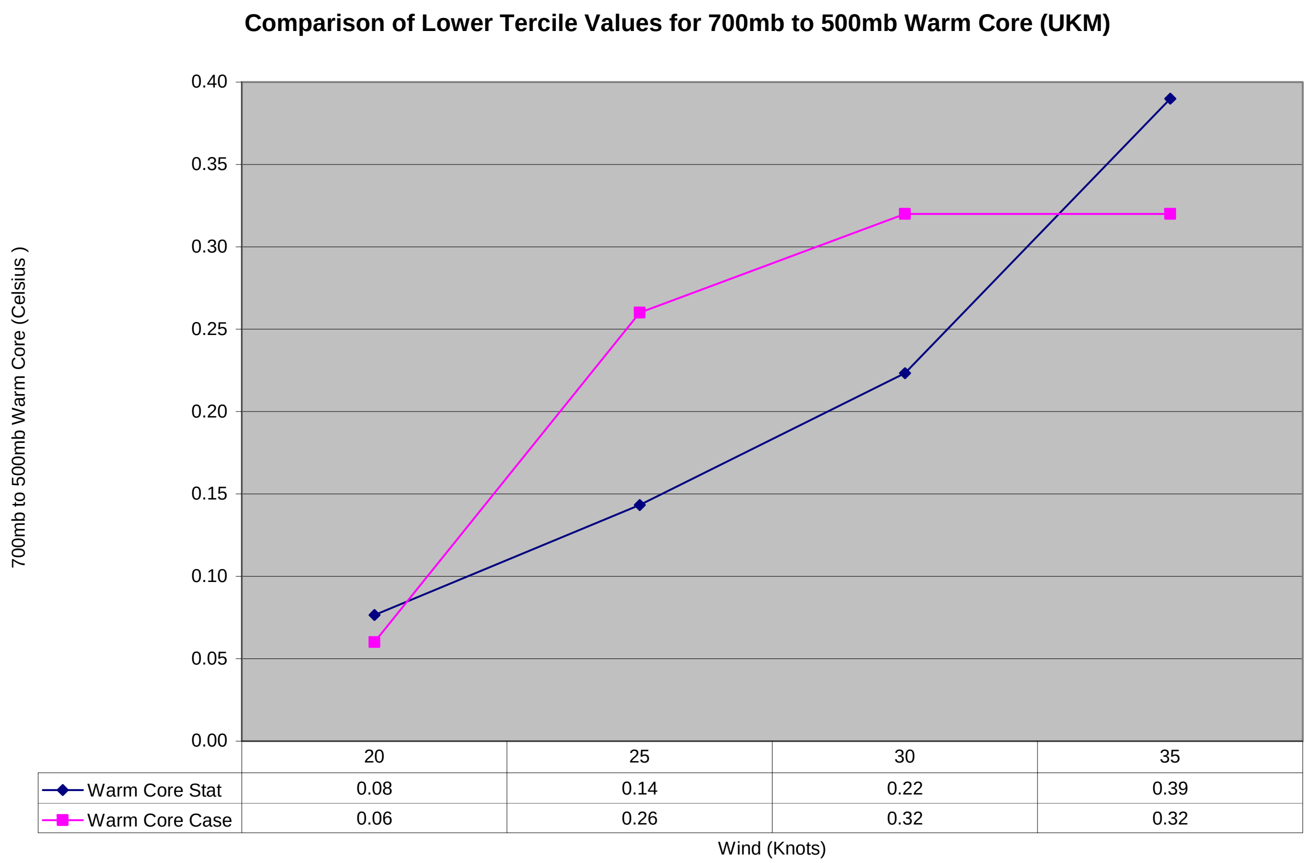
Figure D-4. GFS LTV for case and statistical methods of calculating 700 - 500 mb warm core for 20, 25, 30, and 35 kt



	20	25	30	35
Statistical Layer Vorticity	2.52	2.95	4.06	5.07
Case Layer Vorticity	2.09	3.17	4.37	5.53

Knots

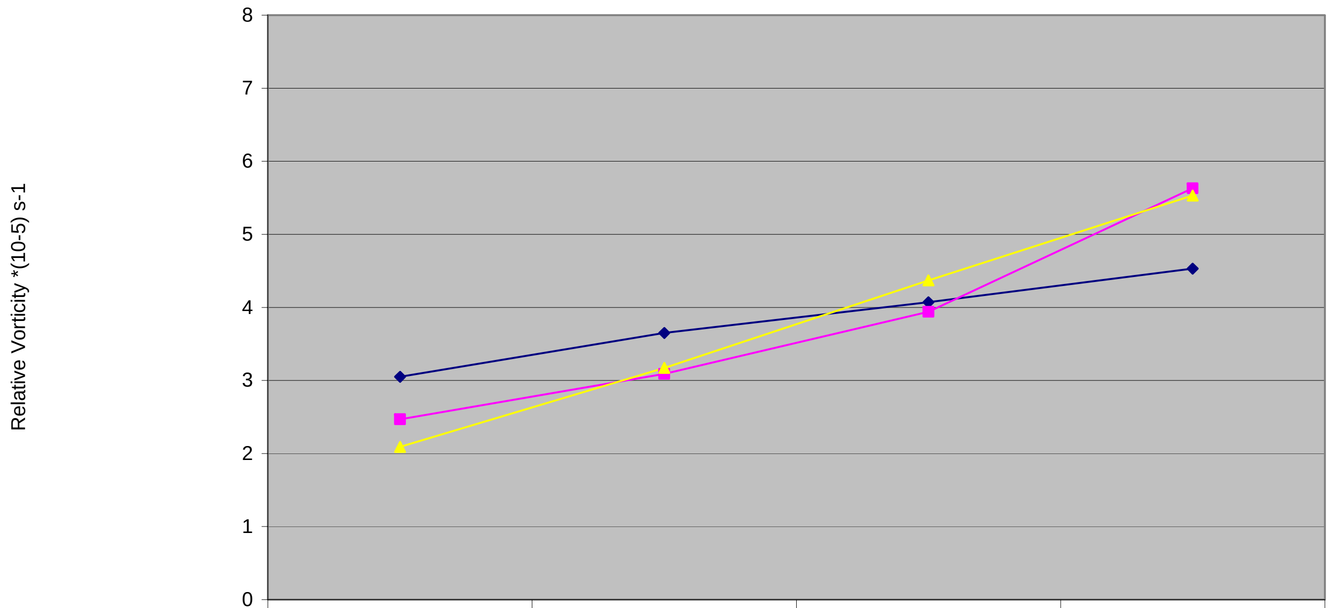
Figure D-5. UKMO LTV for case and statistical methods of calculating 850 - 500 mb relative vorticity for 20, 25, 30, and 35 kt



	20	25	30	35
Warm Core Stat	0.08	0.14	0.22	0.39
Warm Core Case	0.06	0.26	0.32	0.32

Figure D-6. UKMO LTV for case and statistical methods of calculating 700 - 500 mb warm core for 20, 25, 30, and 35 kt

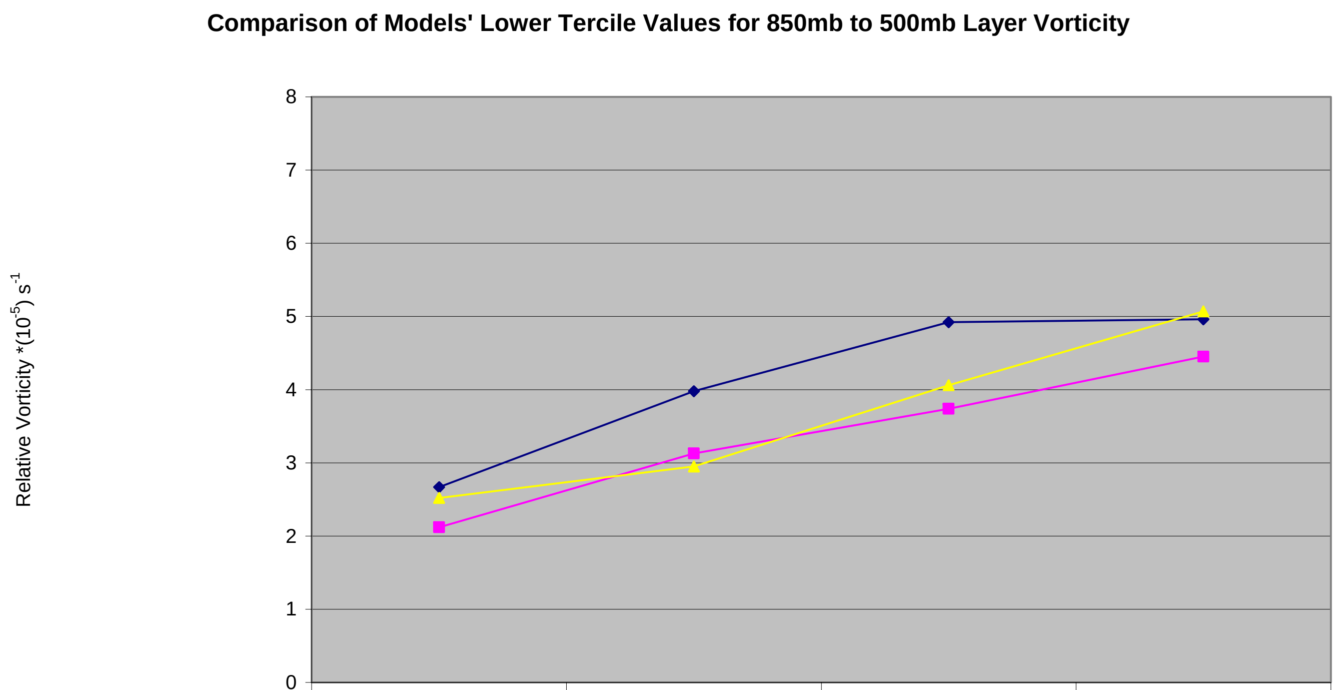
Comparison of Models' Lower Tercile Values for 850mb to 500mb Layer Vorticity



	20	25	30	35
GFS Case Vorticity	3.05	3.65	4.07	4.53
NGP Case Vorticity	2.47	3.09	3.94	5.63
UKM Case Vorticity	2.09	3.17	4.37	5.53

Knots

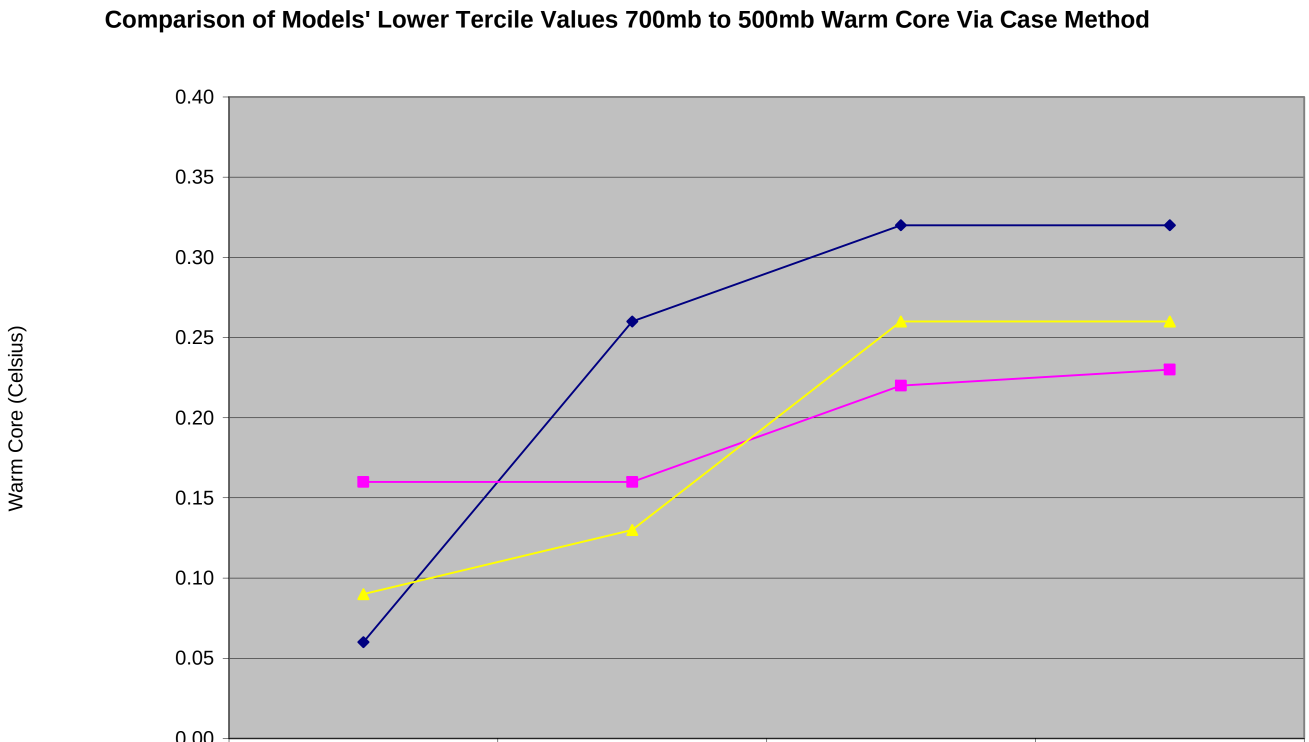
Figure D-7. Comparison of all models LTV for 850 - 500 mb relative vorticity at 20, 25, 30, and 35 kt using case method



	20	25	30	35
GFS Statistical Vorticity	2.67	3.98	4.92	4.96
NGP Statistical Vorticity	2.12	3.13	3.74	4.45
UKM Statistical Vorticity	2.52	2.95	4.06	5.07

Knots

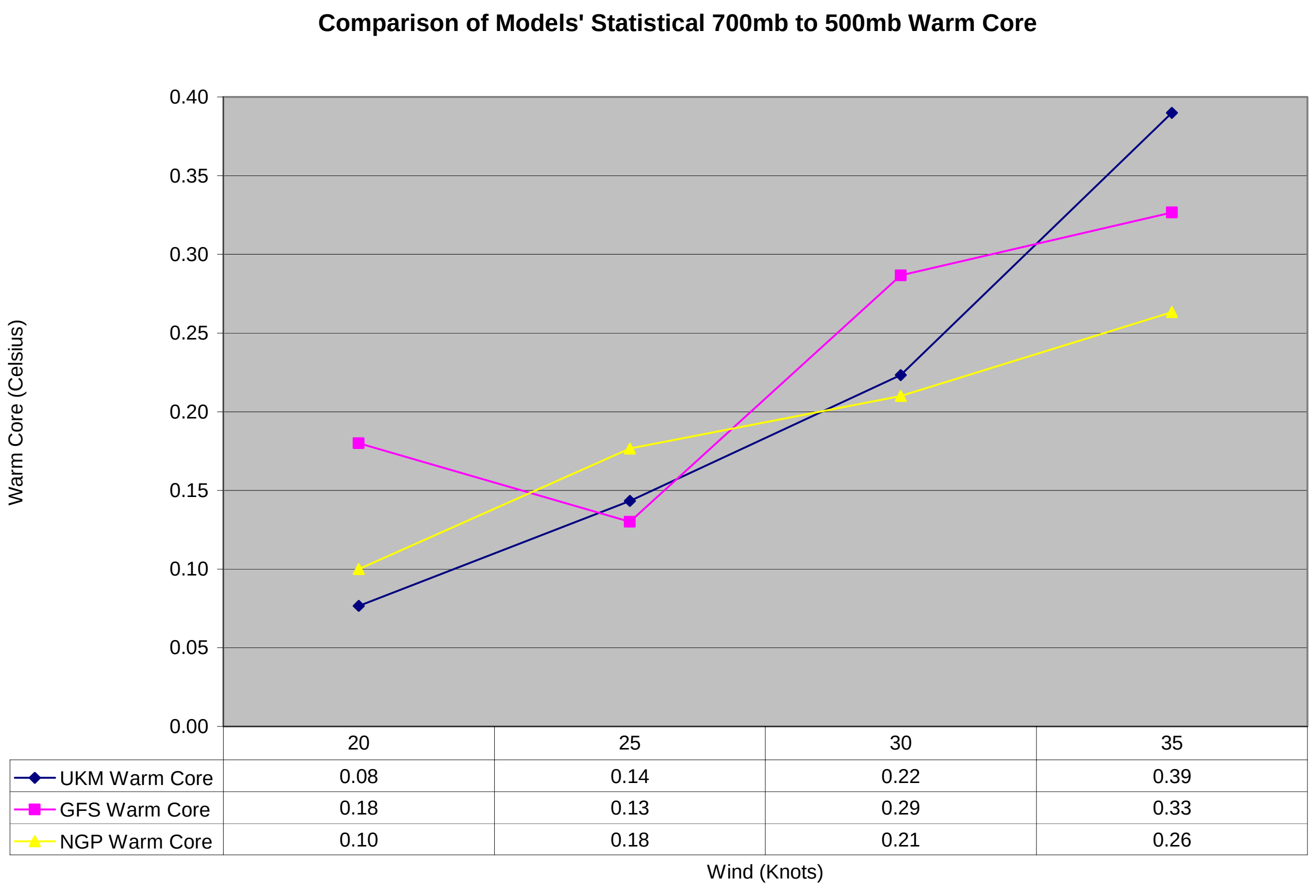
Figure D-8. Comparison of all models LTV for 850 - 500 mb relative vorticity at 20, 25, 30, and 35 kt using statistical method



	20	25	30	35
UKM Warm Core	0.06	0.26	0.32	0.32
GFS Warm Core	0.16	0.16	0.22	0.23
NGP Warm Core	0.09	0.13	0.26	0.26

Wind (Knot)

Figure D-9. Comparison of all models LTV for 700 - 500 mb warm core at 20, 25, 30, and 35 kt using case method



	20	25	30	35
UKM Warm Core	0.08	0.14	0.22	0.39
GFS Warm Core	0.18	0.13	0.29	0.33
NGP Warm Core	0.10	0.18	0.21	0.26

Figure D-10. Comparison of all models LTV for 700 - 500 mb warm core at 20, 25, 30, and 35 kt using statistical method

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Blake, E. S., 2007: Tropical Cyclone Report: Hurricane Humberto (AL092007) 12-14 September 2007. [Available at: http://www.nhc.noaa.gov/ms-word/TCR-AL092007_Humberto.doc] (Accessed February 2008).

Cowan, C. G., 2006: Objectively determined model-derived parameters associated with forecasts of tropical cyclone formation. M.S. Thesis, Department of Meteorology, Naval Postgraduate School, Monterey, CA, 121 pp.

Elsberry, R. L., and D. R. Frill, 1980: Statistical post-processing of dynamical tropical cyclone model track forecasts. *Monthly Weather Review*, **108**, 1219-1225.

Elsberry, R. L., 2003: Tropical cyclone formation. Tropical Meteorology (MR3252) Course Notes, Chapter 2, Naval Postgraduate School, Monterey, CA, 93 pp.

Harr, P. A., 2006: Objective and automated assessment of operational global forecast model predictions of tropical cyclone formation and life cycle. Final Report, Joint Hurricane Testbed Project. [Available at http://www.nhc.noaa.gov/jht/03-05_proj.shtml] (Accessed April 2008).

Peak, J. E., and R. L. Elsberry, 1982: A simplified statistical post-processing technique for adjusting tropical cyclone tracks. *Papers in Meteorological Research*, **5,** 1-14.

Ulses, G. A., 1998: Statistical post-processing of NOGAPS tropical cyclone track forecasts. M. S. Thesis, Department of Meteorology, Naval Postgraduate School, Monterey, CA, 77 pp.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Captain Stephen Chesser
   Naval Postgraduate School
   Monterey, California

4. Professor Patrick Harr
   Naval Postgraduate School
   Monterey, California

5. Professor Russell Elsberry
   Naval Postgraduate School
   Monterey, California

6. Director
   Joint Typhoon Warning Center
   Pearl Harbor, Hawaii

7. Director
   National Hurricane Center
   Miami, Florida

8. Commander
   17th Operational Weather Squadron
   Hickam AFB, Hawaii

9. Commander
   45th Weather Squadron
   Patrick AFB, Florida

10. Air Force Weather Technical Library (AFWTL) at the Air Force Combat Climatology Center (AFCCC) AFCCC/DOO
    Asheville, North Carolina